Naval Training Systems Center

(12)

AD-A169 962

NAVTRASYSCEN 85-C-0044-2

SIMULATOR DESIGN AND INSTRUCTIONAL FEATURES FOR CARRIER LANDING: A FIELD TRANSFER STUDY

D. P. Westra, G. Lintern,
D. J. Sheppard, K. E. Thomley,
R. Mauk, D. C. Wightman,
and W. S. Chambers

12

NAVTRASYSCEN 85-C-0044-2

SIMULATOR DESIGN AND INSTRUCTIONAL FEATURES FOR CARRIER LANDING: A FIELD TRANSFER STUDY

D. P. Westra, G. Lintern,
D. J. Sheppard, K. E. Thomley,
R. Mauk, D. C. Wightman,
and W. S. Chambers

QUALITY INSPECTED 1

Submitted by:

Essex Corporation
1040 Woodcock Road
Orlando, Florida 32803

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Prepared for:

Naval Training Systems Center
Orlando, Florida 32813

Contract N61339-85-C-0044

Interim Final Report for the Period
26 April 1985 - 25 April 1988

DOD DISTRIBUTION STATEMENT

Approved for public release;
distribution unlimited.

NAVTRASYSCEN 85-C-0044-2

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE:

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S):

5. MONITORING ORGANIZATION REPORT NUMBER(S): NAVTRASYSCEN 85-C-0044-2

6a. NAME OF PERFORMING ORGANIZATION: Essex Corporation

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: Naval Training Systems Center

6c. ADDRESS (City, State and ZIP Code): 1040 Woodcock Road, Suite 227, Orlando, FL 32803

7b. ADDRESS (City, State and ZIP Code): Orlando, FL 32813

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Naval Air Systems Command

8b. OFFICE SYMBOL (If applicable):

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N61339-85-C-0044

8c. ADDRESS (City, State and ZIP Code): Washington, DC 20361

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
N637-33N	SP-01	0785-2P3	4780

11. TITLE (Include Security Classification): (See Continuation Page)

12. PERSONAL AUTHOR(S): D. P. Westra, G. Lintern, D. J. Sheppard, K. E. Thomley, R. Mauk, D. C. Wightman, and W. S. Chambers

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Final	FROM 4/85 TO 4/88	18 June 1986	86

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	08	
14	02	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Carrier landing, Transfer of training, Field of view, Scene detail, Incremental transfer, Backward chaining, Visual Technology Research Simulator (VTRS)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A transfer-of-training experiment was conducted as the culmination of the carrier landing behavioral research program at the Visual Technology Research Simulator (VTRS) at the Naval Training Systems Center (NAVTRASYSCEN) in Orlando, Florida. The results of this experiment provide information on the design and use of simulators for training the aircraft carrier landing task. The results also provide input on design issues for the Navy's new T-45 training system (T-45TS). Two visual display variables and two simulator training variables were selected for inclusion in this experiment: scene detail (day contrasted with night); field of view (wide versus narrow); approach type (circling, modified straight-in or segmented); and number of simulator trials (20, 40, or 60). A total of 72 student pilots were trained on the VTRS prior to going through the Field Carrier Landing Practice (FCLP) phase of their pilot training program. The performance of these students at FCLP was contrasted with that of a group of 54 students who did not receive simulator training. Results show that students trained in the

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: Dennis C. Wightman

22b. TELEPHONE NUMBER (Include Area Code): (305 646-5130

22c. OFFICE SYMBOL: N732

11. (Title): Simulator Design and Instructional Features for Carrier Landing: A Field Transfer Study (UNCLASSIFIED)

19. Abstract (Continued)

simulator performed better at FCLP than students ~~in the~~ control group. There was no transfer advantage for those trained with a daytime high-detail scene compared to those trained with a lower cost nighttime low-detail scene. There was ~~also~~ no transfer advantage for those trained with a wide field of view compared to those trained with the lower cost narrow field-of-view scene. Transfer performance was better for the students who had 40 or 60 simulator trials than for ~~the~~ students who had 20 simulator trials. The pilots who trained with a segmented approach schedule did as well or better on transfer to FCLP than those training with the modified straight-in approach schedule or all circling approaches. The segmented approach schedule is recommended as it involves the least time in the simulator.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF TABLES (continued)

SECTION I

INTRODUCTION

Carrier landing, as a task that is both difficult and unique to Navy operations, was the first task selected for intensive study at the Navy's Visual Technology Research Simulator (VTRS). This work on carrier landing started in 1978 when VTRS became operational as a research facility. It emerged as a good candidate for early study primarily because its acquisition consumes a considerable amount of training time, and because it remains a difficult and potentially dangerous phase of flight operations throughout a pilot's career. Furthermore, it appears that carrier landing skills can be taught in a simulator designed for that purpose (Brictson and Burger, 1976).

An effective training simulator for carrier landings should be particularly valuable because the current method of preliminary carrier landing training, to practice approaches to a field site that is specially prepared to simulate features of a carrier landing deck, does not permit practice on some crucial characteristics of the task. For example, the field carrier cannot simulate deck movements found at sea. These movements add to task difficulty and pose significant learning challenges to Student Naval Aviators (SNAs). The presence of air turbulence from the carrier island and the absence of ground effect are other features of carrier operations that are not represented accurately at Field Carrier Landing Practice (FCLP), and that adds to the difficulties encountered by SNAs during their early carrier approaches.

RESEARCH PLAN

The goals of the behavioral research program at the VTRS are to determine simulator design requirements and instructional features for teaching Navy flight tasks. The overall research plan is to examine selected tasks in a series of within-simulator studies and culminates with a study that involves transfer to the airplane. The research reported here is the final (field) transfer experiment that followed the series of within-simulator studies (Collyer, Ricard, Anderson, Westra, and Perry, 1980; Westra, Simon, Collyer, and Chambers, 1981; Hughes, Lintern, Wightman, Brooks, and Singleton, 1981; Westra, 1982; Wightman, 1983; Lintern, Kaul, and Collyer, 1984). It was aimed at identifying optimum design and instructional features for a simulator to teach carrier landings. In addition, this experiment was necessary to determine the ultimate value of tackling major simulator design research problems in the manner currently being employed at the VTRS.

A sequential research plan, progressing from within-simulator to field studies, was developed because a simulator-to-airplane transfer study is both expensive and difficult to conduct properly. The first major step in the sequential research plan was a nontraining or performance study (Westra, 1982). Although this effort provided limited information about training effectiveness for Student Naval Aviators, it did serve to validate experimental manipulations and performance measures. In addition, these data were used as a basis for excluding some experimental manipulations from later training studies. That procedure was founded on the assumption that a variable that does not directly affect performance is unlikely to produce any worthwhile differential effect in subsequent transfer to a standard condition.

Skilled pilots were used as subjects in our performance study so that the results do have some implications for simulators that are used for skill maintenance and transition training. Conditions shown to help pilot performance in the simulator may be considered desirable for skill maintenance simulators, although data from this type of study do not indicate whether there will be any subsequent enhancement of flight performance.

The bulk of within-simulator research at the VTRS has been in the form of quasi-transfer studies. In these studies, a variety of simulator conditions were used to train independent groups of subjects. After predetermined periods of training, subjects were transferred to another simulator condition (the criterion condition) that was as similar to the aircraft as possible. A control group trained and tested on the criterion simulator configuration was also included. This procedure can be contrasted to a true transfer study in which subjects transfer from the training to the operational device. Quasi-transfer refers to a situation in which subjects are tested in the training device, but on a criterion configuration.

Quasi-transfer studies were used to examine variables that may affect training. Those that survived the performance studies, and others that pre-experimental work or other research had suggested would have a worthwhile effect, were tested in this phase. Quasi-transfer studies were first proposed as a means of screening variables for subsequent transfer studies. They provided an economical method of screening the variables of interest in transfer and thereby enabled more effective use of resources available for the transfer studies.

PERFORMANCE EXPERIMENTS. In one of the first experiments at the VTRS, Westra, Simon, Collyer, and Chambers (1982) investigated the effects of 11 simulator-design factors on the performance of experienced pilots. The factors and levels are shown in Table 1. This experiment represented our first attempt to apply the economical multifactor methodology proposed by Simon (1973, 1977).

TABLE 1. EXPERIMENTAL FACTORS FOR A VTRS CARRIER LANDING NON-TRANSFER EXPERIMENT (WESTRA ET AL. (1982)

	LEVELS	
FACTOR	"low"	"high"
FLOLS	Projected CIG	Projected light model
Field of view		
Vertical	-27 to +9 deg	-30 to +50 deg
Horizontal	+/- 24 deg	+/- 80 deg
TV line rate	525	1025
Engine lags	7.5 Hz update	30 Hz update
Ship detail	Night point light	Day solid surface
Visual-system lag	217 msec	117 msec
Seascape	Homogeneous gray	Wave pattern
Brightness		
Ship	0.40 fL	2.90 fL
Sea	0.04 fL	0.50 fL
Sky	0.02 fL	0.16 fL
Platform motion	Fixed base	Six degrees of freedom
Ship type	CIG	Camera/Model board
G-seat	Off	30 pneumatic bellows
Turbulence	None	Close to maximum flyable

Pilots: Eight fleet pilots, experienced in carrier landing.

The only substantial effect on critical measures of task outcome quality came from a comparison of two methods of modelling the Fresnel Lens Optical Landing System (FLOLS). Glideslope tracking performance with a computer-generated FLOLS was better than with a projection from an incandescent-light-source model. This effect was thought to result from the size difference of the two FLOLS simulations. The computer-generated FLOLS had been modeled at larger than scale to overcome limitations in the resolution of the line-scan projector system.

Two other factors had effects which, although smaller in overall impact, were considered potentially important. Approach lineup was better with the day scene than with the night scene, and the shorter visual system lags resulted in less roll variability during the approach. The other factors had negligible effects. The display and simulator factors had been varied over a wide range of interest that represented expensive versus inexpensive simulator options. Thus, the small to null effects resulting from variation of equipment factors in this experiment suggest that the simulator performance of experienced pilots on the carrier landing task is not enhanced substantially by high levels of fidelity.

QUASI-TRANSFER EXPERIMENTS. The information obtained from the performance study aided the design of a subsequent quasi-transfer experiment with pilots who had no previous carrier landing experience (Westra, 1982). In general, if a factor effect was considered practically negligible, the factor was either not studied further or it was combined with others. Thus, the model-board image-generation system was not used again, and the g-seat, TV-line rate, and engine lag factors were not tested further. Elements of scene brightness and seascape detail were incorporated with ship detail into a new scene detail factor. The optical FLOLS was also dropped from further study since it had resulted in poorer performance, even though it was the more expensive of the FLOLS display methods. However, FLOLS size, which was believed to be primarily responsible for the effect, was studied in a later quasi-transfer experiment (Sheppard, 1985).

The factors tested in Westra's (1982) quasi-transfer experiment are summarized in Table 2. Field of view, which had been investigated only for straight-in approaches, was tested in combination with approach type (straight-in or circling). Platform motion was retained for the quasi-transfer experiment primarily because of its high cost implications. Visual system lag was not included because of a concern that it would interfere with the effects of other variables. FLOLS rate cuing (the addition of descent rate information to the glideslope displacement information normally provided by the FLOLS), which had been tested in a performance study (Lintern, Kaul, and Collyer, 1984), was also included as a factor in the quasi-transfer experiment.

TABLE 2. EXPERIMENTAL FACTORS FOR A VTRS WITHIN-SIMULATOR CARRIER LANDING TRANSFER EXPERIMENT (WESTRA, 1982)

FACTORS	LEVELS "low"	LEVELS "high"
Field of view		
Vertical	-27 to + 9 degrees	-30 to +50 degrees*
Horizontal	+/-24 degrees	+/-80 degrees*
Ship detail	Night point light	Day solid surface*
Platform motion	Fixed base	Six degrees*
Approach type	Modified straight-in	Circling*
FLOLS rate cuing	Command	None*
Turbulence	None	Close to maximum flyable
Pilot type	Air Force T-38	Navy P-3C

* Indicates setting for the transfer test configuration. Turbulence was set at an intermediate level setting for the transfer test.

Although small and transient, there were glideslope tracking and lineup advantages in transfer from training with the wide field of view and high scene detail conditions. However, platform motion and FLOLS rate cuing had no differential transfer effects. Approach type did have a substantial effect on lineup, with better transfer performance resulting from training with the straight-in approach. These findings indicated that only field of view, scene detail, and approach type should be tested in a subsequent simulator-to-airplane transfer study.

Data from other laboratories on FOV and scene detail effects have been mixed. A wide FOV appeared to help simulator performances during turns to final and during glideslope tracking (Kraft, Anderson, and Elworth, 1980). However, runway scene detail can modify these effects in that, with some measures at least, performance with a narrow FOV was better than with a wide FOV for low scene detail on the runway. A quasi-transfer study by Collyer, Ricard, Anderson, Westra, and Perry (1980) showed strong FOV performance effects during carrier landing training, but no effect of these training differences on transfer to a criterion simulator condition.

In the Kraft et al. (1980) study, and another by Buckland, Monroe, and Mehrer (1980), scene detail affected landing performances in the simulator. In contrast, Martin and Cataneo (1980) found no effects of simulator training with different levels of scene detail on conventional landing performance in an airplane. The scene detail issue is thrown into further confusion by the work of Brictson and Burger (1976). They showed that prior training with a night carrier display helped night landings but not day landings. This result is puzzling in that the crucial sources of information from outside the cockpit, those being Fresnel Lens Optical Landing System and deck lineup indications, are similarly represented for both day and night landings. Of course, much environmental information is lacking in night landings, but the use that carrier pilots make of that information is not obvious. Taken as a whole, these data indicate that further research on the transfer effects of variations in FOV and scene detail is warranted.

Other quasi-transfer experiments were undertaken at the VTRS to examine instructional features. The most promising result was obtained by Wightman (1983) in a test of a backward-chaining procedure for teaching simulated carrier landings. His experimental subjects were taught carrier approaches in a series in which early trials were started at 2000 feet behind the carrier, and later trials at 4000 feet, and then 6000 feet behind the carrier. This procedure was more effective than whole-task training in which subjects flew all their trials from the 6000-foot mark. The result was perceived as consistent with the advantage shown for straight-in approaches in the Westra (1982) quasi-transfer study. Thus, an approach-type factor combining features of the manipulations tested in both of these studies was developed for the transfer study.

One further important result from the quasi-transfer research was that FLOLS size did not affect acquisition of the carrier landing task (Sheppard, 1985). This had become an important issue because the FLOLS in the VTRS is represented on a larger scale than it is at the carrier. Limitations in the resolution of the out-of-cockpit visual display had forced the larger representation. A true-scale representation, although tested, was impractical because small but critical elements of the display flickered excessively as they crossed raster lines. Engineering redesign would have corrected the problem, but the oversize representation was a more convenient and less expensive solution. The Sheppard study indicated that simulator training with an oversize FLOLS would have no adverse effect on transfer to a normal-size FLOLS in the field.

Several other training manipulations were also examined. Hughes, Lintern, Wightman, Brooks, and Singleton (1981) tested a procedure in which the simulated aircraft approach was frozen on the approach to the carrier whenever a student made a serious error. After the instructor discussed the errors with them, students were either released from freeze to continue their approach or were reset to an optimum position on the glideslope and then released. Wightman (1983) reduced the lags between throttle movement and simulator response during training. Sheppard (1985) fixed his students at a constant (1500 feet) distance from the carrier to allow intensive practice with control responses required for glideslope tracking and also examined the use of descent rate cuing in training. All of these training manipulations were reasonably effective, but were not generally better than the conventional procedures and, therefore, were not included in the transfer experiment.

In summary, two equipment factors, those being field of view and ship detail, were included in the transfer experiment on the basis of the results from the within-simulator research at the VTRS. That research also supported the inclusion of a backward-chaining type of instructional factor, referred to here as approach type. In addition, training time was included as an experimental factor so that incremental transfer effectiveness (Bickley, 1980; Orlansky, 1982; Povenmire and Roscoe, 1971, 1973) could be estimated.

PERFORMANCE MEASUREMENT. Our considerable research experience with the carrier landing task has helped us establish a viable performance measurement approach for the simulator and that same approach seemed preferable in the field. Nevertheless, it proved impractical to gather all measures in the field that had been possible in the simulator. A substantial independent effort to prepare for the transfer experiment identified LSO ratings of approaches as a measure that could be obtained with relative ease and that had strong face validity (Isley and Spears, 1982). However, it also became apparent that LSOs often used their ratings for motivational rather than for assessment purposes; an observation that had been made previously by Brictson and Burger (1976). In some cases the ratings appeared to be more heavily influenced by individual LSO criteria than by student ability.

Further developmental work for performance measurement was undertaken by McCauley and Cotton (1982). A laser system, designated HYTAL for Hybrid Terminal Assist Landing, was identified as being available for the experiment and able to measure altitude and lineup errors in the final approach to touchdown. As measures of altitude and lineup error had dominated the analyses of simulator data, the HYTAL system was considered suitable for the experiment. The loss of Angle-of-Attack (AOA) error data was regrettable, particularly in view of the emphasis that LSOs place on AOA control, but McCauley and Cotton (1982) could not identify a system to measure AOA that could be acquired with the resources available for this experiment.

Statistical power was considered as an important issue. Waag (1980) has noted that many transfer investigations of simulator design issues have not had sufficient power to demonstrate real effects. The use of statistical power to estimate the number of subjects needed in a transfer-of-training experiment has been discussed by Lintern, Thomley, Nelson, and Roscoe (1984). Our requirement in this carrier landing transfer study, to test for differential advantage from various simulator conditions in contrast to merely comparing simulator to no-simulator training, was judged to stress power requirements. Preliminary analyses, using VTRS quasi-transfer data as a guide, together with considerations of achieving a balanced factorial design for the four factors of interest, indicated that 72 experimental (simulator) subjects and a smaller number of control (no-simulator) subjects would be satisfactory.

SECTION II

METHOD

A transfer-of-training design was used to study the effects of two simulator design factors (field of view and ship detail) and two instructional factors (approach type and simulator training time) on carrier landing training. Navy undergraduate pilots visited the VTRS for simulator training immediately prior to their normally scheduled field carrier landing practice which is used as a workup for initial carrier landing qualification. Seventy-two experimental and eight control (no-simulator training) subjects visited the VTRS facility. Field and carrier data were collected on an additional 46 control subjects who did not visit the VTRS.

APPARATUS

SIMULATOR. The Visual Technology Research Simulator (VTRS), described elsewhere by Collyer and Chambers (1978), has a fully instrumented T-2C Navy jet trainer cockpit, T-2C flight dynamics, a six degree-of-freedom synergistic motion platform, a 32-element g-seat, a wide-angle visual system that can project computer-generated color images, and an instructor/ operator control station. The motion system and g-seat were not used in this experiment.

VISUAL SYSTEM. The background scene is displayed by a 1025-line raster system, subtending 50 degrees above to 30 degrees below the pilots' eye level, and 80 degrees to the left and right sides of the cockpit's longitudinal axis. The carrier image, which was a representation of the USS Forrestal, is overlaid on the background with a 1025-line target projector. A carrier wake and FLOLS are also displayed with the target projector. Both daytime and nighttime carrier images are available (Figures 1 and 2).

Average delay between control inputs and generation of the corresponding visual scene is approximately 117 msec. Calculation of new aircraft coordinates requires 50 msec, while calculation of the coordinates for the visual scene corresponding to the viewpoint for the new aircraft coordinates requires approximately 50 msec. Generation of the new scene requires 17 msec. An updated visual scene can be displayed every 33 msec.

The sky brightness for the day scene was 0.85 fL (foot-Lambert) and the seascape brightness was 0.6 fL. The brightest area of the day carrier was 4.0 fL. Except for the horizon, there were no background features represented in either the sky or sea. The night background luminance was 0.04 fL. The

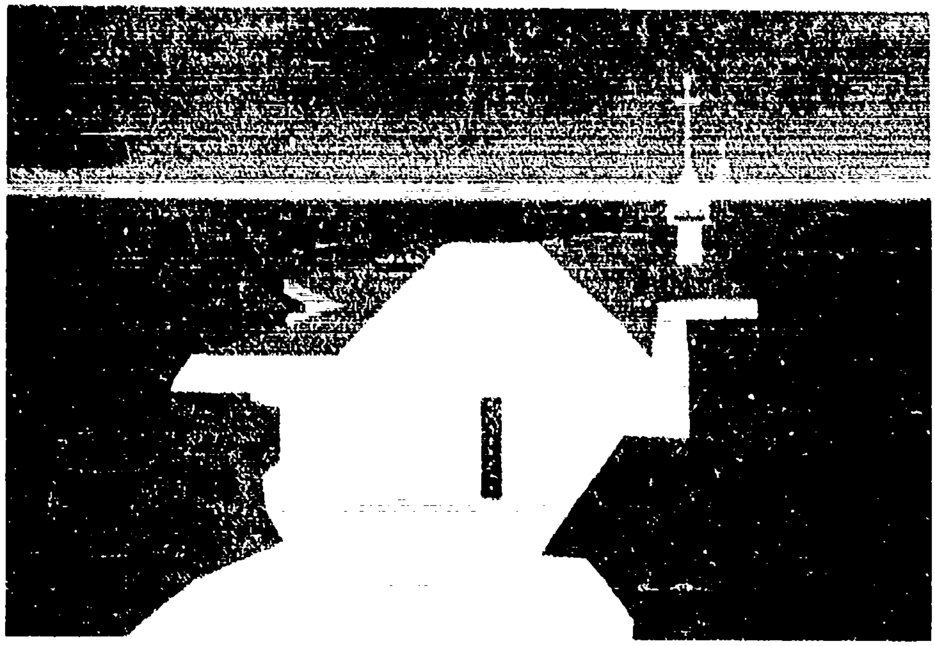

Figure 1. Simulated Day Carrier Scene

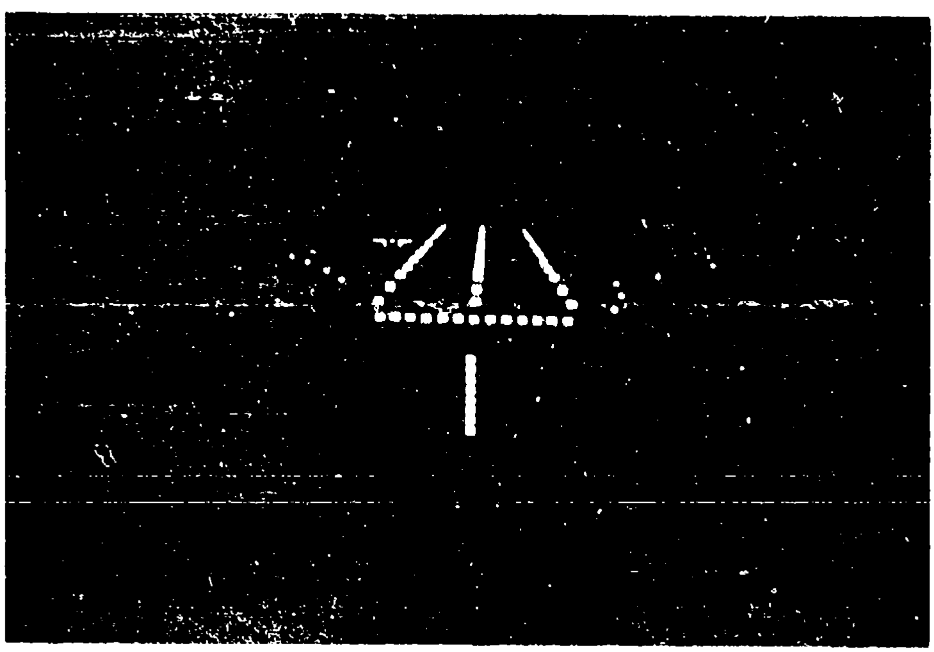

Figure 2. Simulated Night Carrier Scene

horizon was visible but the seascape was not. The night carrier appeared as lights of 0.8 fL brightness outlining the landing deck and other features.

FRESNEL LENS OPTICAL LANDING SYSTEM. The FLOLS and its use are described in Appendix A. The simulated FLOLS is shown with the simulated carrier in Figures 1 and 2. To prevent some of its smaller elements from shimmering and disappearing temporarily as they crossed raster lines, the simulated FLOLS was enlarged by a factor of 4.5 when the distance behind the ramp was greater than 2250 feet. From 2250 feet its size was linearly reduced until it attained 1.5 times its normal size at 750 feet. It remained that size throughout the remainder of the approach. The FLOLS was set for 3.25 degree glideslope.

INSTRUCTOR/OPERATOR STATION. An experimenter and a Landing Signal Officer (LSO) sat at the instructor/Operator Station (IOS) and were able to communicate with the student pilot from that position. Two color monitors, one displaying the background image and the other displaying the target image, provided a general perspective of what the pilot viewed in the simulator.

Two graphics monitors provided the LSO with feedback on student pilots' performance. One display was a real-time view of major cockpit instruments and the other display presented a time history of control and performance measures plotted from a distance of 6000 feet from the carrier to touchdown. Further description of the IOS and graphic displays are included in Appendix B.

FIELD APPARATUS. Flight trials with the T-2C jet trainer were undertaken first at Goliad Field, Texas, and then on the carrier USS Lexington during operations in the Gulf of Mexico. These flights were a normal component of the Navy basic jet training program which culminates with Carrier Qualification (CQ) aboard the USS Lexington. Four FCLP sites in the vicinity of Goliad Field were used at one time or another during the period of data collection for the experiment. Different sites were used depending on weather conditions and convenience. A FLOLS of the type simulated in the VTRS is available both at FCLP and at the USS Lexington.

A laser tracking system (HYTAL) was placed near the active runway during FCLP. This system recorded altitude and lateral deviations from glideslope in the final approach to touchdown (McCauley and Cotton, 1982). An optical glass retroflector was mounted in the landing light housing of each experimental aircraft to return the laser signal. A detailed description of the HYTAL system is given in Lintern, Mauk, and Cotton (in press), and a less detailed description is provided in Appendix C.

PROCEDURE

Approximately 60% of the students from each intermediate class of training squadron VT-26 at NAS Chase Field,[1] Texas, were selected to visit Orlando. Although this selection was intended to be random, it was made by Navy personnel at the training squadron. Nonrandom factors, such as delayed progress of individual students in the course due to poor weather or aircraft mechanical problems, may have influenced the selection to some extent. Experimenters monitored the selection of available students throughout the experiment, and as far as could be determined, no critical biases were introduced. The Orlando visit was made generally on the weekend before the commencement of FCLP. Eight to 12 SNAs, and two to three Landing Signal Officers (LSOs), visited VTRS every six weeks. Data collection extended from October 1983 to August 1984.

The students who did not visit the VTRS were classified as Texas-control subjects. Originally, it was planned that 16 members of each class to visit Orlando would be randomly selected as additional control subjects and would not fly the simulator. However, due to lack of availability, only eight pilots were assigned to this category. This number was considered too small for comparison purposes and these VTRS controls were combined with the other controls who did not visit VTRS.

The students who were to be trained with one of the experimental simulator conditions were briefed on the simulator by VTRS experimenters. They were then briefed on the task requirements and on their simulator training condition by VT-26 Landing Signal Officers who had accompanied them to Orlando. Each pilot had two minutes of preliminary flight time in the simulator prior to experimental trials. Simulator training was conducted in 10-trial sessions with a minimum of one-half hour rest between sessions. The number of sessions for each pilot was dependent on the number of training trials in their experimental condition. Pilots who had 20 or 40 trials completed their training in one day. Pilots who had 60 trials were trained over a two-day period. Each pilot was assigned a VT-26 LSO who monitored their simulator trials throughout training from the instructor/ operator station. LSOs gave instructional advice during and after a trial, much as is done at FCLP and at CQ.

---

[1]The VT-26 squadron is located at NAS Chase Field, but FCLP is normally conducted at nearby Goliad Field. Several LSOs and students fly the T-2C aircraft from Chase to Goliad daily. The remaining students travel to Goliad each morning by bus. A new group of students take over the aircraft during refueling between "events" or "flights."

Following their simulator training, the experimental subjects and the VTRS control subjects rejoined their classmates at NAS Chase Field to continue their undergraduate flight training. The FCLP portion of the course was generally scheduled over 10 consecutive weekdays at Goliad Field. Students generally flew one "flight" each day, consisting of eight FLOLS approaches, to a runway that was marked near the approach end to represent a carrier landing deck. Instructors flew with the students on the first two flights and typically controlled the aircraft for the first flight and half of the second flight. Students flew solo on their remaining eight FCLP flights. Inclement weather, delays for individual students earlier in the course, and aircraft problems could disrupt this schedule so that students occasionally missed flights, or started their FCLP late, and made up that time by flying two events on one or more of the succeeding days.

A major influence on FCLP schedules was the availability of the USS *Lexington*. FCLP was scheduled to start approximately 12 days before the *Lexington* was available, and LSOs went to some trouble to ensure that their students had completed the required number of flights (10) in time to meet that schedule. The students flew to the *Lexington* soon after completing their FCLP and made several approaches and landings at the ship over a two-day period.

SIMULATOR FLIGHT TASKS

All approaches were initialized with the simulator in its landing configuration. Those from 3000 and 6000 feet were initialized with the simulated aircraft on glideslope and lineup: 349.5 degrees heading, 103 knots airspeed, 83% power, and a 500 fpm descent rate. The modified straight-in approaches were initialized with the simulated aircraft in straight-and-level flight, 15 degrees to the port (left) of centerline (see Figure 3), heading 18.5 degrees, 400 feet altitude, 104 knots airspeed, and 86% power. In starting from this position, students were required to fly forward at this altitude and heading to intercept the centerline and glideslope, and then to commence their landing approach. The circling approaches were initialized with the simulated aircraft in straight-and-level flight and 4421 feet off the port beam of the simulated ship, heading 170 degrees, 606 feet altitude, 96 knots airspeed, and 86% power. From this position students were to undertake a descending left turn to intercept the centerline and glideslope for their landing approach.

SIMULATOR FACTORS AND LEVELS

FIELD OF VIEW. The high level field of view was a 160-degree horizontal by 80-degree vertical display (Singer-Link, 1977) which is costly and is representative of that currently

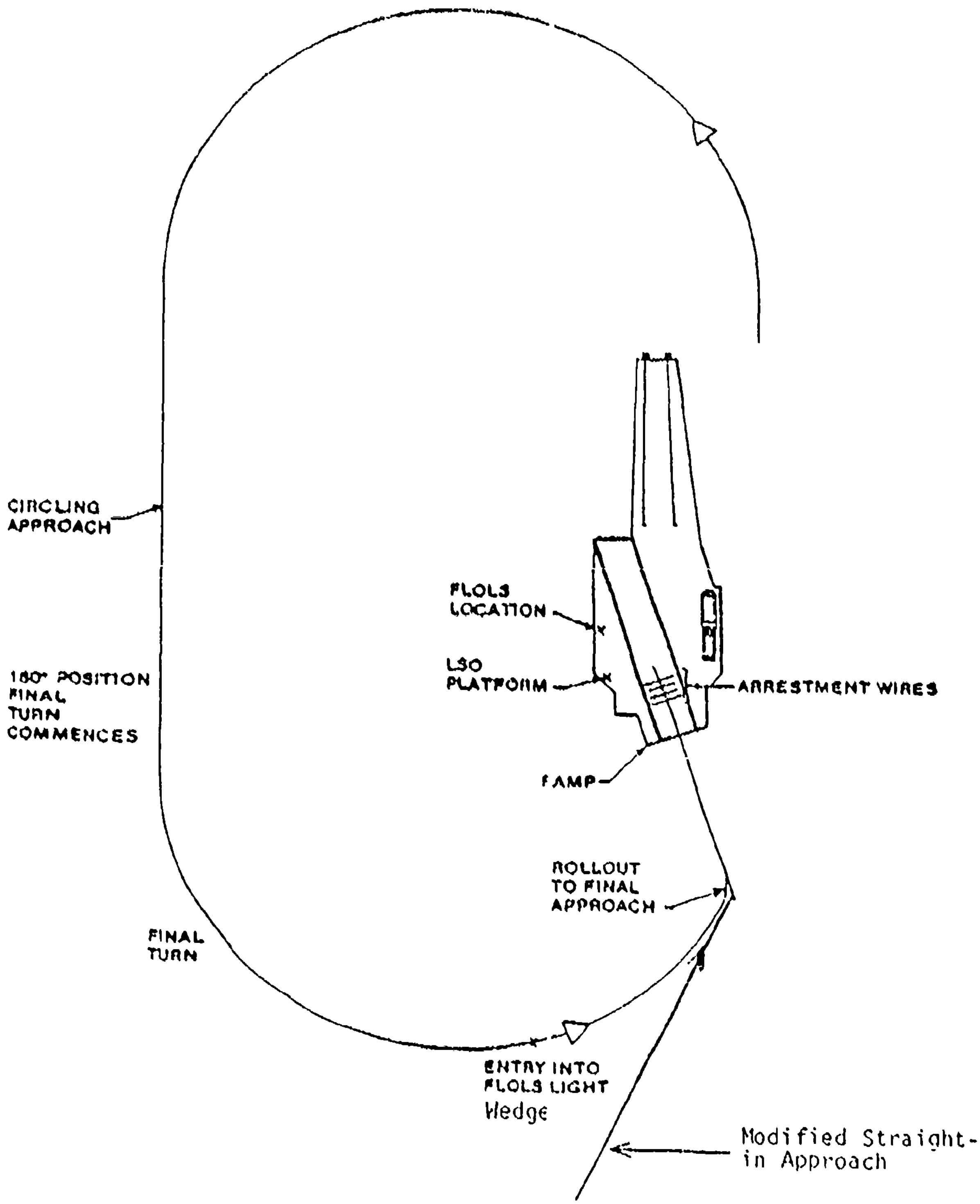

Figure 3. Overhead View of Typical Daytime Circling Carrier Landing Pattern and Modified Straight-in Approaches (not drawn to scale).

available for carrier landing training only on multitask trainers such as the 2B35 and the F-14 Wide Angle Visual System (WAVS). The low level field of view was plus or minus 24 degrees horizontally by -27 degrees to +9 degrees vertically, which is representative of the lower cost Night Carrier Landing Trainers (NCLTs) used for F-4, F-14, A-6, A-7, and S-3 transition training. Figure 4 gives a graphical depiction of the alternative fields of view.

SCENE DETAIL. The high level of scene detail was represented by a daytime solid model CIG (General Electric, 1979) carrier whose surfaces were defined by 985 edges. The daytime scene included wake, light blue sky, and a uniform dark blue seascape below a well defined horizon. This level of detail was approximately representative of that available from daytime CIG systems costing several million dollars, such as the 2B35 trainer, although displayed at higher resolution than available in the 2B35.

The low level of scene detail was represented by an image of a night point-light CIG carrier consisting of 137 lights. It contained all deck outline, runway, centerline, and drop lights. The background was dark with a visible horizon. This display is representative of a night CIG system costing less than a million dollars and used on several Navy NCLTs.

NUMBER OF TRAINING TRIALS. Three levels of VTRS training time were included so that incremental transfer effectiveness could be estimated. Pilots were trained at VTRS for a period of 20, 40, or 60 training trials. However, the amount of simulator time within each session was dependent on the approach-type condition (described below) that was assigned. In addition, the control group who flew no VTRS trials technically provided a fourth level of this factor.

APPROACH TYPE. Three levels of approach type were varied in the experiment. One group of pilots flew all <u>circling</u> approaches while the other two groups performed a backward-chaining sequence of approaches. The <u>modified straight-in</u> group flew the first 75% of their approaches from the modified straight-in start position, and the last 25% of their approaches from the circling start position. The group flying the <u>segmented</u> approach schedule flew the first 25% of their approaches straight-in from 3000 ft, the next 25% straight-in from 6000 ft., the next 25% as modified straight-in approaches, and the last 25% as circling approaches. The initial conditions for each approach were described earlier under the heading "Simulator Flight Tasks." The circling approach took approximately 90 seconds to complete, while the modified straight-in approach took approximately 60 seconds, the 6000 ft. approach 40 seconds, and the 3000 ft. approach 20 seconds. For an equal number of trials, the modified straight-in group had 25% less actual simulator time than the circling group, while the segmented group had 42% less time than the circling

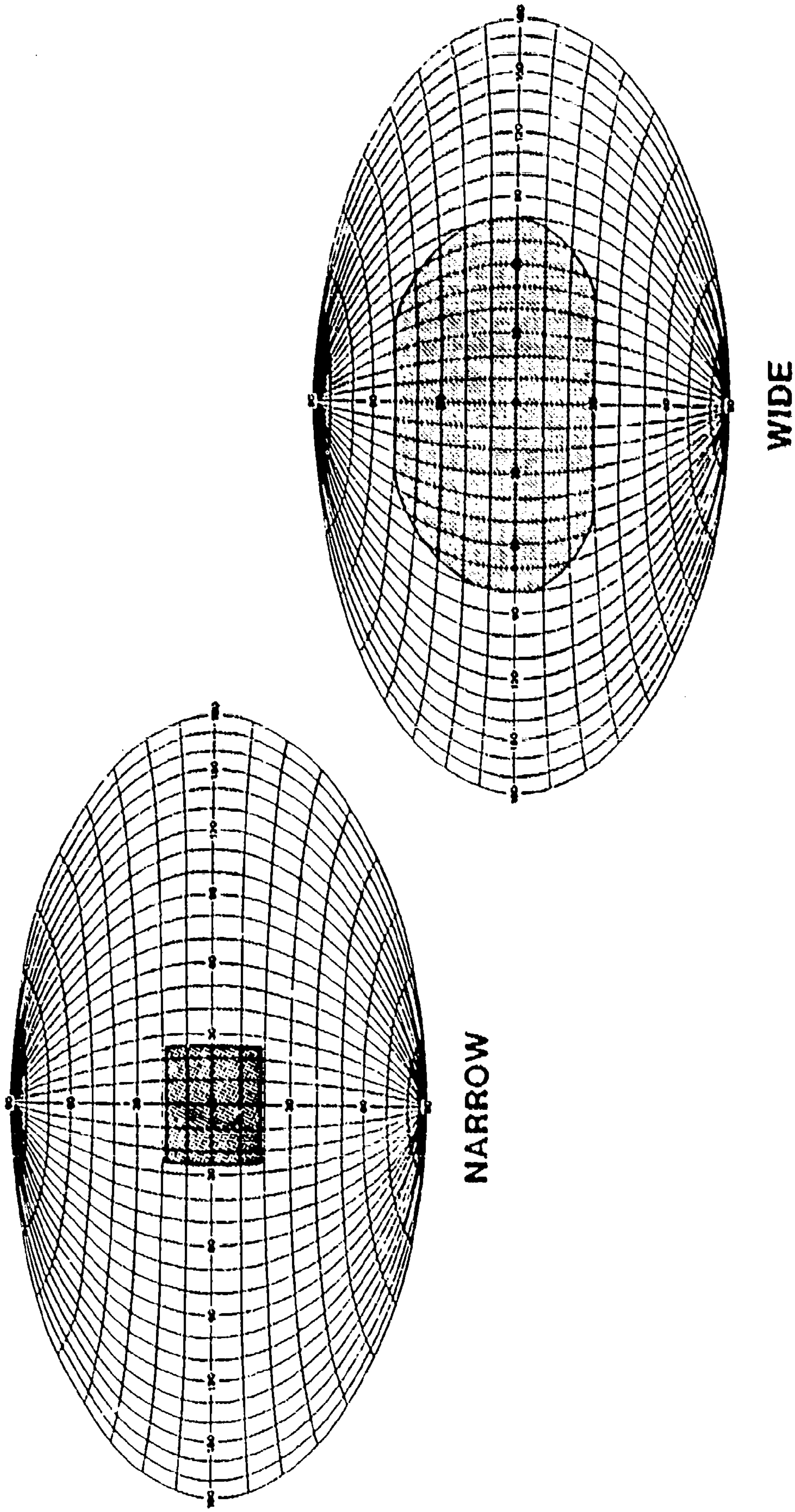

Figure 4. Alternative Field of View Display Plots for Carrier Landing

group. Table 3 gives a summary of the simulator conditions used in the experiment.

TABLE 3. SIMULATOR FACTORS AND CONSTANTS

FACTORS	LEVELS		
Field of View			
Vertical	-27 to +9°	-	-30 to 50°
Horizontal	± 24°	-	± 80°
Ship Detail	Night point light	-	Day solid surface
No. of Trials	20	40	60
Approach Type	Segmented: 25% of trials each with straight-in from 3000 ft, straight-in from 6000 ft, 15° offset left from 8000 ft, circling	Modified straight-in: 75% of trials with with 15° offset left from 8000 ft, 25% circling	All circling approaches

CONSTANTS

Wind speed	5 knots
Wind direction	169.5° (directly down the landing deck)
Ship speed	stationary
Ship heading	360°
Glideslope	3.25°

SIMULATOR PERFORMANCE MEASUREMENT

Parameters of aircraft position and attitude were sampled within the simulator at 30 Hz and were used to derive altitude and lineup error scores from the desired approach path and deviations from desired AOA (15 units). Root Mean Square (RMS) error, mean algebraic error, and variability around those means were calculated for the three performance dimensions over four segments of the final 6000 feet of the approach. These four segments were 6000 ft. to 500 ft. down the ramp, 4500 ft. to 2000 ft., 2000 ft. to 500 ft., and 1000 ft. to the ramp. Time-on-target (TOT) scores were also computed for these segments for time within desired limits in the lineup and vertical dimensions during the approach. Tolerance limits were set at ±0.3 degrees for the glideslope (roughly plus or minus one "ball") and +0.75 degrees for lineup. Pilots flying within these limits will generally be flying safe, high quality approaches without LSO intervention. Distance down the deck, distance from the centerline, and descent rate were measured at touchdown, and the Landing Performance Score (LPS) (Brictson, Burger, and Wulfeck, 1973) was calculated. The LPS is a score assigned to each pass, ranging from 1.0 (technique wave-off) to 6.0 (#3 wire trap).

FIELD PERFORMANCE MEASUREMENT

Glidepath tracking scores for FCLP were measured by the HYTAL system at Goliad Field (see Appendix C). Loss of HYTAL data was occasioned by equipment breakdowns, runway changes due to wind shifts that often found the experimenters still relocating the equipment as the first trials were flown to the new runway, and public holidays or poor weather that forced events to be flown at Chase Field rather than Goliad Field. LSO ratings were, however, available for all FCLP and CQ events.

The HYTAL system sampled aircraft range, azimuth, and elevation positions at 20 Hz at the field. This raw data was later reduced to a 5 Hz sample rate on permanent storage tape and then used to derive summary scores for the approaches. Root Mean Square (RMS) error scores and time-on-target (TOT) scores over the segments 3100 ft. to 1100 ft. from touchdown, 1100 ft. to 100 ft., and 3100 ft. to 100 ft. were used as the primary indicators of approach performance. The independent RMS error components of mean algebraic error and variability about these means were also computed over the segments indicated. Relative position at touchdown was also extracted from this data, but apparent errors in calibration caused inconsistencies in touchdown score transformations.

Within flight trial-to-trial variability scores were also computed as part of the effort to obtain glideslope performance indices that were unaffected by biased glideslope deviation measures. The trial-to-trial variability score was constructed by computing the standard deviation of the mean glideslope deviation values for the approaches within each flight for each pilot. The score was declared "missing" if at least three mean values were not available within a flight. This score is unaffected by glideslope position measurement bias only under the assumption that the bias is constant over a flight. This assumption appeared generally true, although there was also some small bias drift within flights. It is felt that this drift did not pose serious validity threats to the scores. The problem of biased glideslope deviation measures is discussed at length in this report.

LSO GRADE. The LSO in charge assigns a grade to each approach flown during FCLP. This grade is ostensibly a summary indicator of approach and landing quality using the following value designations:

4.0 - reasonable deviations with good corrections
3.0 - reasonable deviations with average corrections
2.5 - bolter
2.0 - below average but safe pass
1.0 - technique wave-off

In practice, a rating of 4.0 is not often given during early FCLP trials and technique wave-offs are fairly rare.

CARRIER QUALIFICATION DATA. Although CQ data are presented in this report, there were reasons to suspect that they would not be sensitive to group differences. Objective performance measures were not available at the ship, and the installation of a glideslope tracking system proved to be impractical. Thus, the only data available from CQ are the LSO grades. In addition, students flew only a small number of approaches at the *Lexington* (generally less than eight) so that the stability of the data was in doubt. CQ data were further compromised by carrier availability. While students flew to a Naval Air Station (NAS) close to the *Lexington* within two days of finishing their FCLP, the time between their arrival at that station and their flights to the *Lexington* varied by several days. During their wait students practiced FLOLS approaches at the field. Thus, the CQ data were judged to provide an insensitive indication of simulator effectiveness because of the subjective nature of the available measures, because of the small number of observations made on each student, and because of the variable amount of practice between FCLP and CQ. Further, the VTRS training was intended to impact FCLP and it was considered very doubtful that transfer would carry over to CQ. This point is discussed in more detail later.

COVARIATE TASK. Ten trials of a psychomotor video game called Air Combat Maneuvering (ACM) were administered to obtain scores that might be used as a covariate in the analysis. ACM is an Atari (TM) video game that has previously shown an association with simulated carrier landing performance (Lintern and Kennedy, 1984). These scores were collected at the VTRS facility for those students who visited Orlando and at NAS Chase Field for those who did not.

EXPERIMENTAL DESIGN

Table 3 lists the four experimental factors manipulated in simulator training, those being field of view (two levels, scene detail (two levels), approach type (three levels), and length of simulator training (three levels). Each of the 72 experimental subjects who visited the VTRS was randomly assigned to one of the experimental conditions. This gave a fully crossed 2X2X3X3 factorial design with two subjects in each cell for the VTRS-trained pilots. In addition, there were 54 control subjects who received no VTRS training. Thus, the experimental design was a factorial experiment with a single control group.

SECTION III

SIMULATOR TRAINING DATA

ANALYSIS OF SIMULATOR TRAINING DATA

The training data represent an unusual situation in which the student aviators flew varying numbers of total training trials. Three levels of training trials were manipulated in the experiment. To equate the number of trials for the repeated-measures ANOVAS and to accurately reflect differences in the amount of training, the data files were constructed as follows:

1) For 20-trial subjects, All trials
2) For 40-trial subjects, Trials 2, 4, 6,...40
3) For 60-trial subjects, Trials 3, 6, 9,...60

Three levels of approach type were also manipulated in the experiment. One group of subjects flew all circling approaches while the other two groups performed a segmented or modified straight-in sequence of trials. Within-simulator assessment of part-task training was possible from comparisons on the fourth quartile of their training trials when all groups were performing the circling task.

The considerable number of performance measures available for analysis of the training data forced selection of those measures that have been validated from previous carrier landing research at VTRS. In addition, log transition was applied to all RMS scores to correct for violations of normality and homogeneity of variance prior to statistical analysis (Levine and Dunlap, 1982). Time in tolerance (TOT), average error, and variability scores were not transformed.

TRAINING PERFORMANCE RESULTS

As argued by Salmoni, Schmidt, and Walter (1984), only transfer data provide an appropriate test of stable differences in learning effects resulting from different training conditions. Nevertheless, several features of the training data were examined briefly for insights they might offer.

Learning effects for performance measures selected for analysis are shown in Table 4. All were significant and substantial. The improvement in performance throughout the training phase was considerable. This evidence was considered as a validation of the performance measures. Lack of a learning effect would have cast doubt on their validity as measures of differential quality of performance. The effects of training time were also significant for all measures with the longer training times being associated with better performance. This result indicated that the selected training times were appropriate for the experiment.

TABLE 4. STATISTICAL SIGNIFICANCE OF THE LEARNING EFFECT IN TRAINING TRIALS FOR SELECTED PERFORMANCE MEASURES

Measure	P
Vertical glideslope, 2000 to 500 ft	
RMS error	0.001
Time in tolerance	0.001
Variable error	0.001
Average error	0.001
Lateral glideslope, 2000 to 500 ft	
RMS error	0.001
Time in tolerance	0.001
Average error	0.001
Angle of attack, 2000 to 500 ft	
RMS error	0.001
Time in tolerance	0.001

Main effects and significant interactions involving FOV, scene detail, and training time are summarized in Table 5. The effects of training time were significant for all measures, with the longer training times being associated with better performance. FOV also had significant training effects but these effects were restricted to the approach lineup performance dimension. Performance across the training trials were better for both lineup TOT and RMS error scores with the wide FOV. These results are shown in Table 5. The scene detail factor did not appear to affect any measures of training performance.

Because of the experimental design, it is possible to view certain contrasts within the training phase for the approach type factor as quasi-transfer (within-simulator) experiments. Performance during the third quartile of trials can be compared for the segmented group and the modified straight-in group to test the effect of the different approach schedules on modified straight-in performance. In the third quartile, both groups were performing modified straight-in approaches, but the segmented group flew 3000 ft. approaches in quartile one trials, and 6000 ft. approaches in quartile two trials, while the modified straight-in group flew modified straight-in approaches during the first two quartiles. The modified

TABLE 5. SUMMARY OF MAIN EFFECTS AND SIGNIFICANT INTERACTIONS OF SCENE DETAIL, FIELD OF VIEW, AND NUMBER OF TRAINING TRIALS: TRAINING DATA

RMS Glideslope Error

Factor	df	F
Scene Detail	1	.00
Field of View	1	.40
No. Trials	2	7.35**

Time on Tolerance Glideslope

Factor	df	F
Scene Detail	1	.14
Field of View	1	.01
No. Trials	2	6.28**

Variability Glideslope Error

Factor	df	F
Scene Detail (SD)	1	.69
Field of View (FV)	1	.00
No. Trials (NT)	2	5.26**
FV X NT	2	4.66*

Average Glideslope Error

Factor	df	F
Scene Detail	1	3.03
Field of View	1	1.07
No. Trials	2	1.07

RMS Lateral Error

Factor	df	F
Scene Detail	1	.82
Field of View	1	7.81**
No. Trials	2	9.81**

Time on Tolerance Lateral

Factor	df	F
Scene Detail	1	1.01
Field of View	1	7.03**
No. Trials	2	6.45**

Average Lateral Error

Factor	df	F
Scene Detail	1	.04
Field of View	1	.21
No. Trials	2	3.80*

RMS Angle of Attack Error

Factor	df	F
Scene Detail	1	.00
Field of View	1	3.08
No. Trials	2	1.60

* $p \leq .05$
** $p \leq .01$

straight-in group did perform better in the third quartile on glideslope tracking than the segmented group. This difference is shown in Table 6 which indicates that the modified straight-in group was within tolerance 58% of the time, compared to 39% for the segmented group. Obviously, the change to modified straight-in approaches was difficult initially for the segmented group. There were no significant differences between these groups on other measures of performance in the third quartile.

TABLE 6. ANALYSIS OF VARIANCE SUMMARY AND MEANS FOR SEGMENTED AND MODIFIED STRAIGHT-IN GROUPS ON THE THIRD QUARTILE OF TRAINING TRIALS: GLIDESLOPE TRACKING (Percent time within ± 0.3 degrees of center)

FACTOR	df	F
Scene Type	1	0.63
Field of View	1	2.25
Number of Trials	2	7.19**
Approach Type	2	18.10**

GROUP	Means
Segmented	39.4%
Modified Straight-in	57.9%

** $p \leq .01$

In the fourth quartile of training trials, all groups were performing circling approaches. This afforded a quasi-transfer type comparison of all three approach type groups on a (within-simulator) circling task criterion condition. There were no differences on the glideslope tracking scores which implies that the disadvantage for the segmented group in the third quartile was transient. Indeed, examination of lineup tracking scores in the fourth quartile indicates that the segmented group performed best. This result is shown in Table 7 and indicates a substantial advantage for the segmented group. The segmented group was within the lineup tolerance limits 64% of the time compared to 44% for the modified straight-in group, and 46% for the circling group. This result is particularly striking when it is noted that the segmented group had not flown any circling approaches until the fourth quartile, while the circling group flew all circling approaches.

TABLE 7. ANALYSIS OF VARIANCE SUMMARY AND MEANS FOR APPROACH TYPE GROUPS ON THE FOURTH QUARTILE OF TRAINING TRIALS: LINEUP TRACKING
(Percent time within ± 0.75 degrees of center)

FACTOR	df	F
Scene Type	1	2.36
Field of View	1	1.39
Number of Trials	2	0.86
Approach Type	2	8.64**.

GROUP	Means
Segmented	64.4%
Modified Straight-in	44.0%
Circling	45.5%

** $p \leq .01$

SECTION IV

ANALYSIS PROCEDURES FOR TRANSFER DATA

Analyzing the flight data proved to be a formidable task in which difficult problems were encountered. Data analysis involved considerable time and effort and yet, despite exploring virtually every possibility, it was eventually determined that no completely justifiable solutions to certain problems were possible. In the interest of timeliness and parsimony only those analyses and results considered to be essentially free of controversy will be presented. It is felt that despite the difficulties, the results presented here accurately and more or less completely describe the outcome of the experiment.

MEASUREMENT ERROR

The most serious problem encountered with the data involved the measurement of vertical glideslope position during the approach. There appeared to be a bias in the measurement of this variable which approached 0.5 degrees in some cases, an amount which far exceeded the 0.1 degree of measurement resolution which was originally specified as allowable for the system. It was concluded that an error in calibration or setting was at fault, but to this date the exact source of the error has not been pinpointed. An attempt was made to pinpoint the problem, but this proved impossible as the HYTAL equipment was packed and returned to China Lake immediately after the experiment and thus was not accessible. The most likely sources of error are in the reference measurement used to calibrate the HYTAL system, calibration error in setting in the FLOLS itself (which was done by the LSO's), electronic drift in the HYTAL system (only a small bias appears possible from this source), drift in the FLOLS setting, or a combination of these.

The bias in measurement could have been statistically adjusted quite simply during data analysis had this bias been constant and of a known value over an identifiable period of time. But the bias eventually was identified as variable on a day-to-day basis within the runway in use for FCLP. Even within this wide range of variability in the bias, a post-hoc adjustment was possible in principle as long as the bias was constant within a known period. However, there was some evidence of measurement drift within a one day period in addition to the bias for that period, and although this drift was not large (not more than 0.1 degree), various attempts at adjusting the data did not appear satisfactory. It was finally concluded that data dependent on the unbiased measure of glideslope vertical position was unrecoverable. Although a great deal of time and effort was expended on these data, to no avail in the case of specific measures, it would seem that the effort was justified when considering the overall cost and importance of the experiment.

Not all vertical glideslope information was lost. Since bias could be considered constant within a trial, the summary measure of variability about the pilot's own mean glideslope was not affected. This measure is considered a useful indicator of pilot control ability and has been used previously to report results from VTRS experiments (Westra and Lintern, 1985). Of course, glideslope measures which summarize performance in terms of overall deviation from the optimal glideslope (including RMS and forms of time-on-target (TOT) scores) were not available as a result of the measurement problems. There was no evidence of a bias or error in the measure of lateral glidepath position. Thus, all summary indicators of lineup performance were not affected by measurement error and were considered "clean." Furthermore, the LSO grades given for the FCLP approaches were available.

MISSING DATA

Despite the fact that over 10,000 approaches were made during FCLP as part of this experiment, there was sufficient missing data to cause serious problems for certain analysis procedures. Data came up missing altogether for a number of reasons. The primary reasons were: 1) Bad weather caused a runway change and the HYTAL technicians were not able to switch the equipment in time to get the data, 2) equipment breakdown, 3) errors in transmitting the data from the video tapes taken at the field to computer tape. The latter was responsible for much of the missing data and apparently was due to recording error, although the exact cause is not known. In addition, there were a significant number of wave-offs given by the LSOs. Most of these were test wave-offs given as a matter of course for training purposes. Usually one test wave-off was given per flight (which consisted of eight approaches) for each pilot. Wave-offs were also given for safety reasons (technique wave-offs) approximately 3% of the time. Whenever a wave-off was given for any reason the objective measures of performance close in to landing were, of course, affected and so this data was not used. Due to all sources more than half of the approaches flown at the field were "missing" in the data summary files for certain measures, even before data editing was undertaken. Although there was nominally still more than sufficient data available for analysis, the real problem was that entire flight means (which were used as the basic data points for analysis procedures) often had to be declared missing.

DATA EDITING

After eliminating all unusable data, the remaining data were plotted and the distributions were examined. Means and standard deviations were taken and this information was used together with the distribution plots to establish cut-off limits for deviant data points. Limits were established on a case by case basis since for some measures there were a few obviously deviant values which could simply be eliminated. In these cases there was most likely a recording error not detected during video tape playback and transmission or a wave-off which had not been recorded properly. In other cases it was not as simple to establish cut-off limits based on examination of the distribution plots without being arbitrary. For these cases a cut-off of six standard deviation units from the mean for the measure was used. These data were probably from more or less complete approaches, but involved performance that went partially "out of control."

FLIGHT MEANS

Flight means for each pilot were computed after data editing was complete and these means were then used as the basic measures of pilot performance for subsequent data analysis and results presentation. Pilots typically flew eight approaches during a flight so that ideally the flight means would represent the average of these eight approaches. However, due to missing data, most pilots had only 3 to 5 valid scores available from a flight. Thus, flight means typically represent the average of 3 to 5 approaches within a flight, but in some cases not enough good data was available to compute a mean. An entire flight mean was considered missing if at least two valid approach scores did not exist within that flight.

LATE TRANSFERS

Unfortunately, not all pilots trained at VTRS went directly to FCLP as planned. This was due to an unforeseen repair delay for the aircraft carrier Lexington which was used for carrier qualifications. FCLP schedules for student pilots are dependent on the availability of an aircraft carrier for carrier qualification which must immediately follow FCLP (i.e., students are not assigned to FCLP unless it is known that a carrier will be available for qualification testing after the completion of FCLP). As a result of this, a number of pilots did not start FCLP for as long as 81 days after completing VTRS training. The actual time delays between VTRS training and commencement of FCLP are summarized in Table 8.

TABLE 8. SUMMARY OF TIME BETWEEN VTRS TRAINING AND THE START OF FCLP

Days Between VTRS and FCLP	Number of Pilots
1 - 8	37
9 - 18	7
34 - 39	8
40 - 47	10
79 - 81	10

Because it can generally be expected that a training benefit will decay with time, the pilots in the experiment were divided into two groups. One group consisted of pilots who transferred to FCLP within eight days of completing VTRS training. The other group consisted of pilots who transferred to FCLP more than eight days after completing FCLP training. In data analysis then, the first significant test performed was a comparison of these two groups. If these groups did not differ, the two groups were combined and compared to the control pilots (not trained at VTRS). If these groups did differ significantly, they were treated separately for purposes of analysis.

ANALYSIS OF TRANSFER DATA

Transfer data were grouped into three major sets prior to performing analysis-of-variance procedures. In the first set data for each performance measure from flights 3 and 4 were combined into a single mean for each pilot. This category (the combination of flights 3 and 4) was labelled "Block 1" and is referred to as such hereafter. These data were considered to be the most sensitive available for detecting the presence of transfer effects and should accordingly be given high priority in terms of judging the size and nature of transfer effects. These data were considered the most likely to show transfer effects because they were the most proximate valid data to the training phase of the experiment. Flights 1 and 2 are, of course, more proximate to the training phase of the experiment, but these passes were flown by instructor pilots in the case of flight 1 and the first half of the passes in flight 2. The second half of the passes in flight 2 were flown by the student pilots, but with an instructor pilot in the plane and typically providing a great deal of verbal assistance. The data from those passes in flight 2 flown by student pilots were in fact analyzed, but results clearly indicated a large instructor effect and so these data were not used to judge transfer results.

Data for the second set, which was analyzed and presented in the Results section, included Block 1 described above and "Block 2" which combined data from flights 5 and 6 into a single mean for each pilot. Data for the third set included Blocks 1 and 2 and "Block 3" which combined data from flights 7 and 8 into a single mean for each pilot. Analyses for data sets 2 and 3 were performed to track the transfer effect (if any) over time. Thus, the analyses of the three major sets provides for an assessment of initial transfer effects and the stability of those effects over time.

PRIMARY CONTRASTS. Tests of significance were made via analysis-of-variance procedures in the following sequence. First, a comparison was made between early and late VTRS transfers as described earlier. If there was no difference between these groups, they were combined into a single VTRS experimental group and compared to the control pilots who did not train at VTRS. Last, differential transfer within the VTRS pilots as a result of the experimental conditions was examined. Within the VTRS trained group, the main effects for the four factors (field of view, scene detail, number of trials, and approach type) were tested along with the two-factor interactions. Of course, if overall transfer for the VTRS group is not present, the interpretation of any within-VTRS group differential transfer effects will be altered.

TEST RESULTS. Analysis-of-variance procedures were used to conduct test of significance. In the case of set 1 (no repeated measures), analysis consisted of standard procedures for a factorial experiment with a single control group (Winer, 1971, p. 468-473). In this procedure the within-group error terms for both the control group and the VTRS group were pooled into a "between subjects" error term and it was this term which the factorial effects were tested against. (Note that the three-way and four way within-VTRS group terms were also pooled into the error term).

Procedures were similar for the analysis of sets 2 and 3 (two blocks and three blocks of transfer data) except that a block (or trials) repeated factor was included in the analyses. In these analyses the block by within-VTRS group factors (main effect and two way interactions) were examined to test for the stability of effects (if any) over time. This procedure made it possible to determine the approximate point at which a transfer differential diminished or decayed. In addition to the within-VTRS group factors, the block by the Control group vs. VTRS group interaction, as well as the interaction involving the block by Early vs. Late VTRS transfers (if a difference existed), was tested.

SECTION V

TRANSFER RESULTS

The transfer results for this experiment are presented separately for the FCLP performance dimensions of glideslope control, lineup control, overall performance (LSO grade), and the carrier qualification scores at the ship following FCLP. Results are presented in the form of flight means for the various experimental groups along with the analysis-of-variance results for the means across the combined flights (blocks) described earlier. The analysis of variance (ANOVA) tables show the sums-of-squares accounted-for by the various terms as well as the percent-variance-accounted-for in the data (eta squared) for the between subject terms. Note that only eta squared values for the between-subject terms are shown for the repeated-measures analyses. These terms sum up to 100% only for the between-subject part of the analyses and are thus not true eta squared values. Eta squared values were reported this way to provide more equitable comparison of results between repeated and non-repeated measure ANOVAs. The tables of means show means for flights 3 through 9, but it should be kept in mind that analyses were performed only on flights 3 through 8. Means shown for flight 9 were often based on considerably less data than the other flights and should be viewed accordingly. A tenth flight was also flown at FCLP, but there was so much missing data from the flight that it was not reported.

The results in general, and for the analysis-of-variance tables in particular, have been condensed in the interest of keeping a reasonable bound on the amount of information presented, particularly in terms of data that might be redundant or overlapping. Thus, each table presenting analyses of variance shows the results for two of the three basic analyses conducted for each FCLP measure. The analysis-of-variance summaries for flights 3 and 4 (combined into a single measure for each pilot as described earlier) gives the best available tests for initial transfer. The results for flights 3 through 8 (combined into 3 blocks for each pilot) then provides tests for effects over the FCLP time period. Analysis-of-variance summaries for data from flights 3 through 6 are not presented but will be discussed where necessary to fully describe the time history of the effect. Additional condensation was done by combining the within-subject block by two-way interactions into a single omnibus term. These interactions were tested individually prior to combining them. However, for the measures presented, there were no significant interactions and so no discussion of the individual terms is necessary.

There were several summary measures available for the performance dimensions of glideslope and lineup control. Within the dimensions, summary measures were available in feet and degrees, over the three approach segments, and in several different transformations (e.g., RMS error and time-on-target transformation for lineup). Much of this information is, of course, redundant and overlapping. In the interests of parsimony, the approach taken was to present complete results for the summary measure which best describes results, and to discuss the other measures as necessary either to more fully describe the result or to provide supporting evidence for a result.

GLIDESLOPE PERFORMANCE

Table 9 gives means for the summary measure of glideslope variability during the approach from 3100 to 100 feet from touchdown. The means show little differences between the early and late VTRS transfer groups and the analysis of variance summaries shown in Table 10 confirm that these groups do not differ significantly, either initially or across flights 3 through 8. Accordingly, the two groups were combined and compared to the control group. Table 10 indicates there was a significant difference between the VTRS-trained pilots and the control pilots. The difference in means can be seen in Table 9. These values indicate that the VTRS trained pilots averaged 0.7 feet less glideslope variability during approaches across flights 3 and 4. This difference appeared to hold up during flight 5 (mean difference of 0.5 ft.), but the analysis of variance for flights 3 through 6 indicated the overall difference across these four flights was not significant. Table 10 shows that there was no overall effect across flights 3 through 8. Thus the initial transfer advantage for the VTRS pilots decayed by about the sixth flight.

None of the VTRS experimental factors appeared to make a significant effect on transfer at FCLP. The closest factor to having an effect was number of training trials. For this factor, the means given in Table 9 suggest that those training with 40 and 60 trials did better at FCLP than those training with 20 trials across flights 3 and 4. Those training with 20 trials differed little from the control group. However, it must be kept in mind that these results were not significant.

The blocks effect is significant across flig ts 3 through 8 as would be expected. However, the effect does not reflect an expected learning trend very well, although most of the block effect does appear to come from improved performance in flights 7 and 8. The lack of a clear-cut learning trend is somewhat puzzling, but it should be kept in mind that this measure is only one aspect of overall performance which the student is trying to improve. There is some evidence to indicate that a student will typically concentrate on one aspect of performance

TABLE 9. FLIGHT MEANS AT FCLP FOR GLIDESLOPE VARIABILITY FROM 3100 to 100 FEET FROM IDEAL TOUCHDOWN

Experimental Groups	Flight 3	4	5	6	7	8	9
VTRS (early)	7.2	6.6	7.1	7.3	7.2	6.5	5.8
VTRS (late)	7.0	6.6	7.8	7.5	6.7	7.5	7.6
All VTRS Pilots	7.1	6.6	7.4	7.4	7.0	7.0	6.7
Control Pilots	7.7	7.3	7.9	7.5	6.5	6.3	5.7
Within VTRS Pilots							
Day Scene	7.3	6.6	7.2	7.4	6.7	6.6	6.6
Night Scene	7.0	6.6	7.6	7.3	7.2	7.3	6.1
Wide FOV	7.0	6.6	7.9	6.9	6.7	6.5	6.2
Narrow FOV	7.3	6.6	7.0	7.9	7.2	7.4	6.5
20 Trials	7.7	7.3	7.5	7.7	6.8	6.9	6.5
40 Trials	7.0	6.2	7.6	7.5	7.1	7.2	6.3
60 Trials	6.8	6.3	7.2	6.9	6.9	6.7	6.3
Segmented Approach	6.8	6.7	7.2	6.9	6.8	7.0	5.6
Modified SI Approach	8.2	6.6	7.6	8.2	7.1	6.9	7.3
Circling Approach	6.6	6.4	7.5	7.1	6.9	7.1	6.4

TABLE 10. ANALYSES OF VARIANCE FOR FCLP GLIDESLOPE VARIABILITY FROM 3100 TO 100 FEET FROM IDEAL TOUCHDOWN

	Flights 3 through 8			Flights 3 and 4		
Source	df	Sums of sq.(%)[1]	F	df	Sums of sq.(%)[1]	F
VTRS (early vs late)[2]	1	7.25	1.3	1	0.14	0.0
VTRS vs Control	1	5.57 (0.9)	1.0	1	17.67 (4.0)	4.4*
Scene Detail (SD)	1	5.20 (0.8)	0.9	1	0.11 (0.0)	0.0
Field of View (FV)	1	5.03 (0.8)	0.9	1	0.34 (0.1)	0.1
No. Trials (NT)	2	14.35 (2.3)	1.2	2	17.35 (3.9)	2.2
Approach Type (AT)	2	14.80 (2.4)	1.3	2	9.44 (2.1)	1.2
SD X FOV	1	0.78 (0.1)	0.1	1	0.77 (0.2)	0.2
SD X NT	2	5.53 (0.9)	0.5	2	5.33 (1.2)	0.7
FV X NT	2	24.62 (4.0)	2.1	2	9.66 (2.2)	1.2
SD X AT	2	5.90 (1.0)	0.5	2	4.98 (1.1)	0.6
FV X AT	2	10.38 (1.8)	0.9	2	1.85 (0.4)	0.2
NY X AT	4	18.46 (3.0)	0.8	4	10.03 (2.3)	0.6
Between Subjects	87	500.75 (81.1)		91	365.97 (82.5)	
Blocks (BL)	2	41.50	4.9**			
BL X (VTRS vs Cont)	2	25.57	3.0			
BL X SD	2	6.61	0.8			
BL X FV	2	8.68	1.0			
BL X NT	4	12.34	0.7			
BL X AT	4	7.27	0.4			
BL X 2-way Int.	26					
Within Subjects	174	732.18				

1 Percent-variance-accounted-for between subjects only.
2 Percent-variance-accounted-for not included in totals.
* $p \leq .05$
** $p \leq .01$

until there is some degree of mastery, then concentrate on another dimension at the temporary expense of the first dimension (Spears, 1985).

OTHER GLIDESLOPE MEASURES. As described earlier, there were serious bias problems with glideslope error data from the laser tracking system. These problems precluded the use of overall error scores such as RMS and time on target (TOT). Under these circumstances, the measure of glideslope variability (in feet) discussed above is considered the best indicator available of glideslope performance. However, other measures which were determined not to be affected by the measurement problem were also examined. Measures of trial-to-trial variability (described earlier under Performance Measurement) were also examined for two approach segments (3100 to 1100 feet and 1100 to 100 feet from touchdown). Results for the trial-to-trial variability measures were generally supportive of the findings presented above although the differences between VTRS and control pilots were not as strong. VTRS pilots performed slightly better than the control pilots in both segments indicating the effect described above was general throughout the approach. It should be noted that the methods used to create the individual pilot trial-to-trial variability measures created additional "missing" flight means which makes these results less powerful than for the measure of within approach glideslope variability.

Glideslope variability in degrees was also analyzed, but inconsistencies due to outliers from the in-close data resulted in a non-conclusive outcome. Within-trial glideslope variability for the approach segments 3100 to 1100 feet and 1100 to 100 feet were also examined, but for this measure it was determined that segmented length was insufficient. Based on examinations of approach flight paths, it appeared that the average period of an average actual flight path was too long, particularly for the 1100 to 100 foot segment. A full period of corrective tracking activity within a defined segment must exist in order to obtain a sensible variability value.

LINEUP PERFORMANCE

The data from the HYTAL tracking system in the lateral dimension appeared to be correct and without measurement bias (other than within the resolution of the system). Therefore, the summary scores of overall error were available and considered the most appropriate indicators of lineup performance. For these data, it was felt that the time-on-target (TOT) summary scores were the best indicators of performance in the lineup dimension. Use of these scores reduced the amount of missing data and alleviated the sometimes subjective task of editing for outliers. Further, the score itself is immediately interpretable in terms of "operationally meaningful" effect magnitude. Tolerance limits were set at

+0.75 and so the TOT measure simply gives the percentage of time within this envelope. Anytime the pilot is within this envelope (and does not have an excessive error rate or change) he is in an "OK" status as far as lineup is concerned.

Table 11 presents the FCLP flight means for the TOT measure over the approach from 3100 to 100 feet from touchdown. These means suggest that the VTRS trained pilots did better at FCLP than the control pilots and the analysis-of-variance summaries presented in Table 12 confirm this. In fact, the test of significance for this comparison was significant across flights 3 through 8 as well as across flights 3 and 4. However, an examination of the means given in Table 9 suggests that the transfer differential has dissipated by the sixth flight. No other effects were significant although there was a suggestion that the groups defined as early and late transfers differed across flights 3 and 4. In this case the early transfer group appeared to be better than the late transfer group as one might expect. Accordingly, the results for the early group alone were examined. However, the results for the four within-VTRS experimental factors showed no significant effects and mean differences did not differ markedly from those shown for the combined early and late transfer groups.

The overall transfer effect for the approach from 3100 to 100 feet from touchdown was further investigated to determine if the transfer advantage for the VTRS pilots occurred in close, at the start and middle, or was relatively constant along the entire approach. To examine this, the lineup TOT score was analyzed for the 3100 to 1100 ft. and the 1100 to 100 ft. segments separately. The means for these segments are presented in Table 13 for the VTRS and control pilots. Table 13 suggests that there was a difference in the 3100 to 1100 ft. segment with VTRS-trained pilots performing better. Analysis of varianace summaries for these data show that the differences in the 3100 to 1100 ft. segment were significant across flights 3 and 4 and across flights 3 through 8. On the other hand, the differences in the 1100 to 100 ft. segment were not significant. Therefore, the transfer effect was restricted to the early part of the approach and not in close. Apparently, the VTRS training improved the ability of pilots to set up and start the approach on the lineup dimension. Although an in-close benefit was not indicated, the improved ability to get a good start and perform better in the middle is obviously beneficial.

OTHER LINEUP MEASURES. RMS error scores for lineup performance were also analyzed and results essentially paralleled those given above. Because the RMS error data involved considerable outlier editing, it is felt the TOT scores most accurately and completely show the outcome of the experiment. Therefore, the RMS error data will not be presented here. The lineup variability and bias components of RMS error were also examined

TABLE 11. FLIGHT MEANS AT FCLP FOR PERCENT TIME WITHIN $\pm 0.75^{\circ}$ OF CENTER LINE FROM 3100 TO 100 FEET FROM IDEAL TOUCHDOWN

Experimental Groups	Flight 3	4	5	6	7	8	9
VTRS (early)	80	86	83	75	75	79	80
VTRS (late)	78	76	75	81	76	83	84
All VTRS Pilots	79	81	79	78	76	81	82
Control Pilots	76	75	72	76	77	82	73
Within VTRS Pilots							
Day Scene	77	79	78	77	75	82	81
Night Scene	82	83	80	80	76	81	86
Wide FOV	77	83	78	77	74	84	85
Narrow FOV	82	79	80	80	77	79	79
20 Trials	76	79	79	80	77	81	83
40 Trials	84	82	77	79	77	83	75
60 Trials	78	84	80	75	73	79	84
Segmented Approach	77	81	79	80	70	72	81
Modified SI Approach	82	79	78	77	79	78	74
Circling Approach	79	83	79	77	80	85	85

TABLE 12. ANALYSES OF VARIANCE FOR FCLP LINEUP TOT SCORES FROM 3100 TO 100 FEET FROM IDEAL TOUCHDOWN

	Flights 3 through 8			Flights 3 and 4		
Source	df	Sums of sq.(%)[1]	F	df	Sums of sq.(%)[1]	F
VTRS (early vs late)[2]	1	0.017	0.4	1	0.061	2.7
VTRS vs Control	1	0.168 (4.5)	4.5*	1	0.097 (3.8)	4.2*
Scene Detail (SD)	1	0.053 (1.4)	1.4	1	0.053 (2.0)	2.3
Field of View (FV)	1	0.009 (0.2)	0.2	1	0.000 (0.0)	0.0
No. Trials (NT)	2	0.042 (1.1)	0.6	2	0.049 (1.9)	1.1
Approach Type (AT)	2	0.013 (0.3)	0.2	2	0.000 (0.0)	0.0
SD X FOV	1	0.007 (0.2)	0.2	1	0.000 (0.0)	0.0
SD X NT	2	0.002 (0.1)	0.0	2	0.002 (0.1)	0.0
FV X NT	2	0.057 (1.5)	0.8	2	0.032 (1.3)	0.7
SD X AT	2	0.037 (1.0)	0.5	2	0.032 (1.3)	0.7
FV X AT	2	0.012 (0.3)	0.2	2	0.010 (0.4)	0.2
NY X AT	4	0.067 (1.8)	0.4	4	0.112 (4.4)	1.2
Between Subjects	87	3.271 (87.5)		91	2.136 (84.7)	
Blocks (BL)	2	.020	0.7			
BL X (VTRS vs Cont)	2	.010	0.3			
BL X SD	2	.030	1.0			
BL X FV	2	.010	0.3			
BL X NT	4	.053	0.9			
BL X AT	4	.039	0.6			
BL X 2-way Int.	26					
Within Subjects	175	2.670				

1 Percent-variance-accounted-for between subjects only.

2 Percent-variance-accounted-for not included in totals.

* $p \leq .05$

TABLE 13. FLIGHT MEANS FOR LINEUP TOT SCORES FOR THE APPROACH SEGMENTS 3100 to 1100 FEET AND 1100 TO 100 FEET FROM TOUCHDOWN

	Flight						
3100 to 1100 ft	3	4	5	6	7	8	9
VTRS Pilots	82	89	82	83	81	88	86
Control Pilots	77	79	77	82	83	77	77
1100 to 100 ft							
VTRS Pilots	76	68	73	69	65	68	75
Control Pilots	74	68	65	64	68	64	66

and indicate that the transfer effect was mostly composed of the variability component. This means that the control pilots were more variable (erratic) in their lineup performance through the 3100 to 1100 ft. segment, not simply biased in one direction or the other compared to the VTRS pilots.

LSO GRADES

Each pilot received a grade from an LSO for each pass at FCLP. This grade is supposed to be an indicator of overall performance during the approach. Previous research (Collyer, Ricard, Anderson, Westra, and Perry, 1980) suggests that the primary components of the grade in order of importance are: 1) touchdown accuracy, 2) approach glideslope control, 3) approach lineup control, and 4) approach angle-of-attack control. The LSO grades represent the most complete data available for FCLP performance and thus represent the most powerful analysis of results (at least in terms of number of data points). However, since the score itself is a composite, it does not necessarily represent the most powerful analysis of components that make up the score. Further, these grades contain error components due to less than perfect within and between rater reliability. The grades are also used for motivational purposes at various points in training and this further decreases objectivity and validity. Nevertheless, they are important scores and should be viewed as such.

Table 14 gives the FCLP flight means for LSO grades for the various experimental groups. Inspection of this table suggests that there are virtually no differences between groups on this measure. The means are very similar, differing by no more than 0.1 unit in most cases. Table 15 gives the analysis of

TABLE 14. FLIGHT MEANS AT FCLP FOR LSO GRADES

Experimental Groups	Flight 3	4	5	6	7	8	9
VTRS (early)	2.36	2.48	2.58	2.66	2.73	2.85	2.83
VTRS (late)	2.34	2.48	2.59	2.59	2.74	2.70	2.85
All VTRS Pilots	2.35	2.48	2.59	2.63	2.74	2.80	2.84
Control Pilots	2.33	2.49	2.59	2.67	2.73	2.74	2.83
Within VTRS Pilots							
Day Scene	2.36	2.49	2.60	2.63	2.67	2.75	2.84
Night Scene	2.35	2.48	2.57	2.63	2.81	2.85	2.85
Wide FOV	2.34	2.42	2.54	2.60	2.72	2.79	2.81
Narrow FOV	2.38	2.55	2.62	2.65	2.75	2.81	2.87
20 Trials	2.36	2.45	2.53	2.61	2.66	2.76	2.86
40 Trials	2.40	2.47	2.60	2.68	2.83	2.82	2.71
60 Trials	2.32	2.54	2.62	2.60	2.75	2.80	2.96
Segmented Approach	2.42	2.54	2.61	2.64	2.76	2.79	2.85
Modified SI Approach	2.32	2.41	2.58	2.54	2..69	2.77	2.80
Circling Approach	2.30	2.51	2.58	2.67	2.78	2.85	2.79

TABLE 15. ANALYSES OF VARIANCE FOR FCLP LSO GRADES

	Flights 3 through 8			Flights 3 and 4		
Source	df	Sums of sq.(%)[1]	F	df	Sums of sq.(%)[1]	F
VTRS (early vs late)[2]	1	0.040	0.4	1	[.000]	0.0
VTRS vs Control	1	0.001 (0.0)	0.0	1	.007 (0.1)	0.1
Scene Detail (SD)	1	0.000 (0.0)	0.0	1	.013 (0.2)	0.2
Field of View (FV)	1	0.129 (1.1)	1.3	1	.018 (1.3)	1.5
No. Trials (NT)	2	0.052 (0.5)	0.3	2	.038 (0.6)	0.3
Approach Type (AT)	2	0.088 (0.8)	0.5	2	.116 (1.7)	1.0
SD X FOV	1	0.044 (0.4)	0.5	1	.012 (0.9)	0.2
SD X NT	2	0.050 (0.4)	0.3	2	.192 (2.9)	1.6
FV X NT	2	0.097 (0.9)	0.5	2	.123 (1.9)	1.0
SD X AT	2	0.178 (1.6)	0.9	2	.079 (1.2)	0.7
FV X AT	2	0.055 (0.5)	0.3	2	.074 (1.1)	0.6
NY X AT	4	0.414 (3.6)	1.1	4	.190 (2.9)	0.8
Between Subjects	103	10.280 (90.2)		103	5.72 (85.9)	
Blocks (BL)	2	1.860	24.7**			
BL X (VTRS vs Cont)	2	0.082	1.1			
BL X SD	2	0.132	1.8			
L X FV	2	0.096	1.3			
BL X NT	4	0.096	0.6			
BL X AT	4	0.111	0.7			
BL X 2-way Int.	26					
Within Subjects	205	7.730				

[1] Percent-variance-accounted-for between subjects only.

[2] Percent-variance-accounted-for not included in totals.

** $p \leq .01$

variance summaries for these scores and confirms that none of the groups differ significantly. There is a strong learning component across groups suggested by the highly significant blocks effect, but the groups themselves do not differ either initially or across flights 3 through 8.

CARRIER QUALIFICATION SCORES

The VTRS training was not designed to impact carrier qualification (CQ) scores but it is still interesting to examine these data. As described earlier, the VTRS training was done with a fixed (non-moving) ship to approximate FCLP conditions rather than the boat for carrier qualification. Further, after approximately 80 FCLP approaches it would be doubtful that any VTRS transfer effect would carry over to CQ. Nevertheless, the means for CQ LSO grades are given in Table 16.

As Table 16 suggests, there appears to be only trivial differences between the groups. Table 17, which gives the analysis of variance summary for these data, confirms that there were no significant differences between any of the groups. Thus, it is concluded that there was no carryover of any transfer effect through FCLP to the boat, at least none that was reflected in the LSO score.

OTHER RESULTS

As described earlier, a great amount of time and effort was expended trying to adjust the bias of glideslope (vertical) position information that was acquired from the HYTAL laser tracking system. Part of that effort involved the construction of a touchdown accuracy score based on glideslope position at touchdown. In fact, the results for this constructed score were presented at at least one briefing (Wightman, Westra, and Lintern, 1985). Although it is felt that the methods used to adjust the raw data were justifiable, it was decided not to present those results here. First, there was still some uncertainty about the accuracy of the constructed score. Second, later more detailed examination of those data revealed some inconsistencies which were difficult to explain. Finally, the results, which showed a large transfer advantage for the VTRS-trained pilots, were inconsistent with other findings. It is believed that the results presented in this report, which show a more modest transfer advantage for the VTRS pilots in both the vertical and lateral dimensions of control during the approach only, accurately and properly (without any controversy) portray the true results of the experiment.

COVARIATES. Several variables were examined as potential covariates for the field carrier landing task. A variable capable of accounting for some significant portion of variance at FCLP could increase the power to detect experimental effects. It could be also useful for predictive purposes as

TABLE 16. CARRIER QUALIFICATION GRADE MEANS FOR THE EXPERIMENTAL GROUPS

Experimental Groups	Means	Experimental Groups	Means
VTRS (early)	2.67	Within VTRS Pilots	
VTRS (late)	2.72	Wide FOV	2.71
All VTRS Pilots	2.70	Narrow FOV	2.68
Control Pilots	2.80		
Within VTRS Pilots		20 Trials	2.71
Day Scene	2.71	40 Trials	2.64
Night Scene	2.60	60 Trials	2.72
		Segemented Approach	2.73
		Modified SI Approach	2.63
		Circling Approach	2.73

TABLE 17. ANALYSIS OF VARIANCE FOR
CARRIER QUALIFICATION GRADES

Source	df	Sums of sq. (%)	F
VTRS (early vs. late)	1	[.049]	0.3
VTRS vs. Control	1	.372 (2.02)	2.2
Scene Detail (SD)	1	.006 (.03)	0.0
Field of View (FV)	1	.020 (.11)	0.1
No. Trials (NT)	2	.123 (.67)	0.4
Approach Type (AT)	2	.123 (.67)	0.4
SD X FOV	1	.135 (.73)	0.8
SD X NT	2	.350 (1.90)	1.0
FV X NT	2	.035 (.19)	.2
SD X AT	2	.098 (.53)	.3
FD X AT	2	.034 (.18)	.1
NT X AT	4	1.08 (5.85)	1.6
Between Subjects	104	16.09 (87.16)	-

well as to provide insight into basic attributes necessary for good carrier landing performance. Previous work had indicated that an air combat manuvering (ACM) video game might be predictive of FCLP performance (Lintern and Kennedy, 1984). In addition to ACM, a crosswind variable (vector component relative to runway multiplied times windspeed), and runway (with code values assigned to the various runways used at FCLP) were examined. The crosswind and runway variables did not prove useful as covariates. They generally did not correlate significantly with the FCLP performance measures which were reported here. During the approach crosswind did show a relationship with the lineup bias scores. However, these measures were not given much weight in assessing the outcome of the experiment.

The 10 ACM trials were categorized into means for the first five and last five trials. These means were significantly correlated (r=0.87) which agrees with previous results showing a high degree of internal reliability and differential stability. However, there were no worthwhile correlations (none greater than r=0.20) between these ACM scores and FCLP LSO grades, approach glideslope scores, or approach lineup scores. One possible reason for these low correlation's may have been the poor reliability of the FCLP scores. Flight-to-flight reliability for LSO grades ranged from r=0.15 to r=0.44 and differential stability (at approximately r=.40) did not appear to be reached until flight 8. Results were similar for the approach summary measures with flight-to-flight reliabilities ranging from r=0.00 to r=0.46. These low reliabilities could be due to several sources, but most likely are due to task difficulty resulting in low within-subject reliability, particularly in the earlier flights. Whatever the reason, the low FCLP flight to flight reliabilities virtually precludes the use of covariates to increase power.

RECYCLED PILOTS. Although no meaningful formal analysis or test of significance is possible, (because of the low numbers involved) it is useful to consider the number of recycles in the experimental groups. A total of seven pilots out of the 126 VTRS and control pilots were recycled through FCLP because of LSO decisions that they were not ready to attempt carrier qualification. Four of these pilots were from the control group and three were from the VTRS group defined as "late" transfers. Table 18 gives a summary of this and indicates that there were no recycles within the VTRS group that transferred to FCLP within eight days of VTRS training. A further look at the pilots who were recycled that were in the VTRS "late" transfer group indicates that these three pilots had delays of 39, 44, and 47 days (6-7 weeks) between VTRS training and FCLP. In other words, there were no recycles among the pilots who trained at VTRS and transferred to FCLP within a reasonable period of time.

TABLE 18. GROUP MEMBERSHIP OF RECYCLED PILOTS

Experimental Group	No. of Recycles
VTRS (early)	0 (0.0%)
VTRS (late)	3 (8.6%)
Control Pilots	4 (7.4%)

SECTION VI

DISCUSSION/RECOMMENDATIONS

This experiment was a major undertaking in terms of time, expense, and logistics. It is remarkable that the effort was carried out and brought to completion with as few problems as there were. It was disappointing to be unable to recover vertical glideslope displacement measures, but enough solid data were obtained to provide an accurate, realistic, and relatively complete summarization of results.

JUDGING THE SIZE OF EFFECTS

In any complex experiment of this kind, with multiple performance dimensions, it is sometimes difficult to judge the operational or "real world" significance of results. In the case of this experiment several statements can be made with certainty, and interpretation of overall effect size must be made based on these and all the other available information. The statements that can be made with certainty regarding transfer effects are: 1) There was a significant advantage for VTRS-trained pilots at FCLP on the measure of approach glideslope vertical variability through flights 3 and 4, after which there was no advantage. 2) There was a significant advantage for VTRS-trained pilots at FCLP on approach lineup control through flights 5 and 6. 3) There were no differences between the VTRS and control groups at FCLP (or subsequent carrier qualification) on LSO grade. 4) There were no significant differences at FCLP for any of the four experimental factors varied at VTRS. With these facts in hand, the discussion that follows will bring to bear other relevant information in an attempt to fully consider and interpret the results.

GLIDESLOPE AND LINEUP APPROACH PERFORMANCE. The size of the transfer effect should be judged from several viewpoints. First, consider the percent-variance-account-for (eta squared) values. In the case of both lineup and glideslope variability, the overall transfer effect (VTRS vs. Control, see Tables 10 and 12) accounted for just over 4% of the variance in these measures across flights 3 and 4. Although it is arbitrary to associate specific eta squared values with "small", "moderate" and "large" effect sizes, one guideline is given by Cohen (1977). He associates eta squared values of 1%, 6% and 11% with small, medium, and large respectively. Based on the authors' own experiences with data of these type, they concur that these values are appropriate. Based on this, then, the effects for the two approach measures can be said to be somewhere between "small" and "moderate"; tending toward moderate.

The duration of the transfer effect should also be taken into account when judging overall size. As previously stated, the effect for one measure lasted approximately through the fourth flight, and through the sixth flight for the other. Including flights 1 and 2 (flown largely by instructors); these durations represent 40% and 60% of the FCLP training period. Again, it seems that the judgment would be for effect sizes somewhere between "small" and "moderate."

LSO GRADES. The transfer effects for glideslope and lineup performance must be balanced against the fact that no transfer differences were found for the LSO grade. From one perspective, this suggests that the overall transfer effect cannot be considered large, which was also suggested in the preceding discussion considering only the glideslope and lineup scores. From another perspective, there is a question as to why at least a trend for a real transfer effect did not show up in the LSO grade. Although it is superficially disturbing that no such trend occurred, an examination of the score elements reveals that the score is inherently insensitive to group differences. Despite the fact that data for this FCLP measure were relatively complete, and thus more powerful from a "numbers" viewpoint than some of the objective measures, there are several other sources of measurement error which tend to weaken the sensitivity of the LSO grade.

As discussed previously, there are sources of measurement error for the LSO grade due to less than perfect between and within scorer reliability. Furthermore, the grade tends to be used at FCLP for motivational purposes which causes validity problems (the score is measuring something besides purely objective approach and landing performance). Also, as discussed previously, the score is a composite of several dimensions of performance. This would tend to weaken its sensitivity if the effect were concentrated in just one or two of the dimensions. In fact, the effect does seem concentrated in the middle of the approach and not touchdown. Yet the touchdown is given the most weight in the LSO grade. But even more importantly, sensitivity of the score is poor because of its structure. In most cases pilots receive either a value of 2 (below average but safe pass) or 3 (reasonable deviations) for their approaches. Values of 4 (reasonable deviations with good corrections) and 1 (wave-off other than practice wave-off) do not occur very often and the value 2.5 (bolter) reflects only touchdown performance when it is given. In one sample of 1198 aproaches, a value of 4 was given 8.3% of the time; 3, 41.7%; 2, 27.5%; and 1, 2.8%. Thus, little discrimination other than in a crude sense is available from the score. The notes in the score sheets describing pilot deviations are of much more value for training purposes than the score itself.

POWER REDUCTION DUE TO PROCEDURE. The reduction in sensitivity to detect effects caused by certain procedural (not measurement) problems must also be considered when judging the transfer effect sizes. These were not problems with experimental design or concept, but rather were inherent to the "real world" field data collection at FCLP. The first problem involved the procedure at FCLP of instructors in the aircraft with the student pilots during the first two flights. The instructor actually flies about three-fourths of the passes during these first two flights and gives intense, one-on-one instruction throughout. Thus, not only is there the time delay until valid measurement can begin in the third flight, but intense additional instruction and practice has been inserted during the interim. This can't help but weaken the ability to measure effects; for example, if a one-flight transfer effect were present there would be no way to measure it. This will be considered further in the discussion of the within-VTRS factors.

The other major problem that weakened the ability to detect transfer effects was the delay in time between VTRS training and the start of FCLP caused by the breakdown of the aircraft carrier used for training. Only a little over half of the VTRS-trained pilots started FCLP within one week of their VTRS training while the others experienced delays of up to 2-1/2 months. An effort was made to separate pilots based on delay time and there was some suggestion that those who transferred early did better at FCLP on lineup performance than those who transferred late. However, the power to examine the within-VTRS factors for this group was weakened both by lower numbers and imbalance in the groupings (there was no control over condition assignments based on early vs. late transfer since this was not anticipated when the experiment began). Thus little additional information was obtained by separating the groups based on early vs. late transfer other than to note that the early transfers did better for some measures.

Taking all this into account, it is not surprising that more or larger effects were not detected. In fact, when considering these issues, more credence is lent to the transfer effects that were detected. It is the authors' opinion then that a transfer advantage of reasonable and moderate size did occur as a result of VTRS training. It is difficult to state a specific value for the transfer effectiveness ratio, but it seems reasonable to state that one or two FCLP flights could be saved by simulator training. But rather than talk about "saving" flight time, the authors would prefer to point out that there were <u>no</u> recycles among the VTRS-trained pilots who transferred to FCLP within a reasonable period of time. It would appear to us that the real value of simulator training for the carrier landing task lies in improved readiness and the reduction of recycles.

In terms of an overall transfer effect, a final point to consider is that VTRS training involved a fixed carrier for maximum transfer to FCLP conditions. However, the ultimate goal of all training is readiness for carrier qualification at sea under operational (moving ship) conditions. Since FCLP cannot provide the moving ship experience, it would seem appropriate to provide additional simulator training with a moving ship near or at the end of FCLP prior to carrier qualification. It would appear to us that such additional training would further improve overall readiness and improve performance at the boat. The issue is an important one, and an experiment designed to test this hypothesis along with the timing of the additional training in the FCLP schedule is recommended.

THE VTRS EXPERIMENTAL FACTORS

The primary purpose of this experiment was to determine whether high cost simulator display options for field of view and scene detail would result in a transfer advantage compared to much lower cost (but operationally reasonable) options. Further, we wished to determine the best way to use the simulator for training in terms of number of trials and schedules of approach. An a priori assumption for meaningful tests of these factors was that overall transfer as a result of VTRS training would occur. As discussed above, this transfer has been documented for certain FCLP measures but not for others. Therefore, it only makes sense to discuss the transfer differences due to VTRS conditions for those measures in which overall transfer was demonstrated.

FIELD OF VIEW. There were no transfer differences at FCLP noted that were due to the field-of-view (FOV) conditions. Based on this information, the recommendation for simulator display design for this task is quite straightforward; the narrow angle display is recommended for the carrier landing task. However, given the problems with FCLP data discussed above, it will also be useful to consider other results. Field of view has been studied previously at VTRS in a performance experiment (Westra, Simon, Collyer, and Chambers, 1981), a quasi-transfer experiment (Westra, 1982), and elsewhere in a quasi-transfer experiment (Collyer, Ricard, Anderson, Westra, and Perry, 1980).

In the Westra et al. (1981) performance study the FOV was tested at roughly the same levels as in this experiment and no differences were found for task outcome scores. There was a difference found for a measure of roll variability with pilots flying under the narrow FOV condition having greater roll variability during approaches. Presumably, this result was simply a function of not having as much horizon with which to judge roll attitude. Overall, the FOV effect was judged as small to marginal. There were also differences in approach performance due to the field-of-view factor during the training phase of the present experiment. Pilots training at VTRS with the wide FOV had less lateral (lineup) error during training than those training with the narrow field of view (see Table 5). In the Collyer et al. (1980) quasi-transfer experiment, pilots were tested under an in-simulator wide FOV criterion condition after training under either a wide or narrow FOV. No differences on the criterion condition were detected for several measures of approach performance, landing performance scores, and the LSO grade. There were substantial differences during training with those training under a wide FOV and circling approaches performing better than those training with a narrow FOV and circling approaches.

Westra (1982) also tested field of view in a quasi-transfer experiment. The FOV levels for training were similar to those tested in the present experiment, and the criterion task had a wide FOV display with full circling approaches. The wide FOV training condition did result in some advantage on the transfer task for final approach quality but not landing accuracy. There was essentially no advantage after eight transfer trials (equivalent to one flight at FCLP). This experiment had both a high degree of control and a good deal of power, and probably represents the best that could be expected from a wide FOV display. This being the case, it is not surprising that no differences were detected in the present experiment where the first valid transfer data started with the third flight. In summary, it would seem that some benefit from a wide FOV could be expected. However, all evidence indicates that this benefit would be fairly small and transient. Therefore, it is our finding that a wide FOV display is not justified for the carrier landing task.

SCENE DETAIL. The scene detail levels which were compared in this experiment did not result in any observable transfer differences at FCLP. Based on these results, it would appear that the higher cost associated with high scene detail as represented in this experiment is not justified for simulating the carrier landing task. Again though, all the information gathered on this variable should be considered. Originally, this somewhat global factor was studied by Westra et al. (1981) as three separate dimensions in a performance experiment. These dimensions were ship detail (day vs. night), seascape and brightness. Lineup performance was affected by ship detail (better performance with the daytime ship) but the seascape and brightness factors had no meaningful effects. In the subsequent quasi-transfer experiment (Westra, 1982), these three dimensions were combined into a scene detail factor with levels defined as in the present experiment. The criterion or transfer task in the quasi-transfer experiment had the high level daytime scene display.

Several transfer effects were noted in the quasi-transfer experiment. First, there was a temporary (one flight) advantage on transfer for the high scene detail for glideslope performance and probably for angle-of-attack performance during the approach. There was also a transfer advantage for the high scene detail on lineup performance but only when straight-in approaches were used during training. There were no transfer benefits noted for touchdown performance. Given the well controlled nature of this quasi-transfer experiment, it would appear then that the maximum benefit that could be expected from a high scene detail display would last about one flight at FCLP. As with field of view, there was little possibility of detecting an effect of this size in the present experiment with no valid transfer data available until the third flight. In summary then, it appears that although some benefits for the high detail scene could be expected, the cost is simply not justified in an operational trainer for the carrier landing task. Simulators with visual displays similar to the displays in the Night Carrier Landing Trainers (NCLTS) currently in use in the fleet would appear to be most cost-effective for training student Naval aviators.

APPROACH TYPE. FCLP data analyses did not reveal any significant differences due to approach type training conditions. Based on this finding there can be no doubt about recommending the segmented approach type schedule for use under operational training conditions. With this schedule the first 25% of the training approaches are flown straight-in from 3000 ft. aft (behind) the carrier. The next 25% of the training trials are flown straight-in from 6000 ft. Then 25% more of the approaches are flown from 8000 ft. behind the carrier and offset 1800 ft. port (left) of the ship's runway centerline. The final 25% of the approaches are flown as circling approaches starting from the abeam position. Thus, if 40 training trials were specified, for example, a student pilot would execute 10 approaches from each of the four starting positions. This schedule saves approximately 42% of the simulator time it takes to run an equal number of all circling approaches and offers (at least) equal training benefit.

Other research has indicated that backward chaining the task actually provides better training in certain performance dimensions than full task (circling) approaches, despite the fact that less time in the simulator is involved. Wightman (1983) compared performance for subjects who trained for a 6000 ft. straight-in whole (simulated) task under backward chaining (approaches starting at 2000 ft., then 4000 ft., then 6000 ft.) and whole task (6000 ft.) conditions. The backward-chaining procedure was more effective in training the whole task, although it must be kept in mind here that the criterion condition was not the full circling task. In another quasi-transfer experiment, Westra (1982) trained pilots for the whole circling task under either all circling approaches or all modified straight-in approaches. The modified straight-in approach group started approximately 12,000 ft. behind the ship in this case (to equate for time with circling approaches), 4150 left of centerline, and heading 18 degrees to the right of the ship's heading. The group training under this conditions did better on line-up performance on the circling transfer task than the group training with circling approaches, despite never flying the full circling task until the transfer test. Although in this case time in the simulator was equated, the results indicated a significant benefit as a result of practicing the part task, with no negative benefits.

Originally, the modified straight-in approach was set up to provide training on certain elements of final turn procedures with the ship in view at all times during the narrow field-of-view display. It now appears that backward chaining provides significant advantages as a training procedure. Therefore, a backward-chaining procedure similar to the one used in the present experiment is recommended for operational use regardless of the display specifications.

NUMBER OF TRIALS. VTRS pilots trained for 20, 40, or 60 trials (approaches) in this experiment. The intent with this factor was to obtain incremental transfer information that could be used to define optimum use of an (operational) trainer in terms of training time. Although there were no significant differences in the FCLP data reported here, there were definite trends established. In fact, in one analysis for flights 3 and 4 for the FCLP glideslope variability score for the VTRS early transfer group only (not presented here), there was a significant ($p < .05$) differential transfer due to number of training trials which supports this trend. This trend can be seen for flights 3 and 4 for the glideslope variability and lineup TOT scores which are presented in Tables 9 and 11. These data suggest that 20 training trials results in transfer no different from the control group, while the groups training at VTRS with 40 and 60 trials did better than the control group. But the 60-trial group generally did no better than the 40-trial group. Clearly then, the recommendation is for 40 training trials in an operational simulator, spaced in sessions of 10 trials each as was the case during VTRS training.

However, in keeping with the notion of using a simulator to reduce recycles and for remedial training, it would be better to think of 40 simulator trials as a base, with more trials given to those who appear to have not mastered essential skills. Further, as discussed previously, it may be desirable to add more simulator trials later or at the end of FCLP with a moving carrier to aid in preparation for carrier qualification.

RECOMMENDATIONS FOR OTHER DESIGN OPTIONS. Before making recommendations for simulator design based on results from previous research, it seems worthwhile to comment on our philosophy regarding cost-effectiveness and fidelity. First of all, fidelity itself should not be a goal for simulator design, but rather training effectiveness. Each design option should then be considered in terms of its cost and its training effectiveness. There are several combinations of cost and training effectiveness which may occur. In one case the cost may be high and the training benefit low (e.g., motion platform). In this case the recommendation is simple; the option obviously is not cost-effective. Other cases may involve options with relatively low cost and low training effectiveness. In these cases, we would recommend the option if it clearly improved fidelity and thus face validity and pilot acceptance. Factors which directly affect the pilot/aircraft response relationship (e.g., lag) bear special consideration since pilots will "learn" to fly whatever the simulation is. We feel that a close proximation of the simulated aircraft response system is important for a valid simulation. Table 19 gives our recommendations for simulator design for the carrier landing task for all design options examined experimentally at VTRS, not including the present experiment. These recommendations are based on the findings together with considerations discussed above.

RECOMMENDATIONS FOR T-45 TRAINING SYSTEM

During the evolution of the VTRS carrier landing research program, a specific need for the results of research became apparent. This need involved design guidance for the T-45 training system (T-45TS). The T-45TS will ultimately replace the T-2C and TA4J aircraft currently used in the Jet Undergraduate Training Program (JUPT). The plan for T-45TS includes extensive use of simulators for training several flight tasks, including carrier landing. The fact that these simulators will be used to teach a number of tasks limits the application of the results somewhat, although the results clearly lead to specifications for the optimum (cost-effective) configuration for training the carrier landing task. Further, cost-effective use of simulators for teaching the carrier landing task has been defined in terms of approach type and number of trials.

The recommendations resulting from the carrier landing research program should be thought of as minimums for the T-45TS simulator designs. The other tasks which are to be taught in the simulators must be studied individually to determine if more costly design options are necessary. In some cases, it may be possible to study certain task dimensions from different tasks which represent common elements, but in general, performance is task specific and thus the individual (distinct) tasks must be considered separately. Given an already defined minimum simulator configuration for one task,

TABLE 19. RECOMMENDATIONS FOR OTHER SIMULATOR DESIGN OPTIONS

Factor	Recommendation	Comments
Platform motion	No motion	No differences found
G-seat	No g-seat	No differances found
FLOLS	CIG	Needs to be greater than actual size
TV Line rate	525	For AOI display with 4:1 zoom
Engine update	30 hz	Although no marked differences were found, the relatively small cost is justified
Visual lag	< 133 msec	Longer lag causes attitude control problems
Ship Type	CIG	The flexibility afforded by CIG displays vastly outweighs any model board advantage
Seascape	Homogeneous background	Wave pattern is not necessary but well defined horizen is
FLOLS rate cuing	Rate cuing for fleet	Does not appear to enhance training for student pilots. Considerably improved performance of experienced pilots
Simulator freeze	"Freezing" during approach not recommended as a training tool	Freeze may contribute to simulator sickness

future research on other tasks should focus on whether additional design features are needed, and other research already conducted should be evaluated from this viewpoint. For example, the VTRS research does suggest that a wider field of view and higher level of scene detail are necessary for training the air-to-ground bombing task than are necessary for the carrier landing task (Lintern, Sheppard, McKenna, Thomley, Wightman, and Chambers, 1985).

SECTION VII

SUMMARY

A field transfer-of-training experiment was conducted to define simulator design requirements and training procedures for the carrier landing task. Two visual display factors and two simulator training factors were included in the experiment. The factors were scene detail (day vs. night), field of view (wide vs. narrow), approach type (segmented, modified straight-in, circling), and number of simulator trials (20, 40, 60). Student Naval Aviators were trained in the Visual Technology Research Simulator in the carrier landing task under the experimental conditions prior to going through the Field Carrier Landing Practice (FCLP) phase of their pilot training program. Performance was then assessed at FCLP which served as the transfer condition. Other student aviators, not pretrained in the VTRS, were used for control comparison.

Pilots who received additional carrier landing training at the VTRS did significantly better at FCLP on measures of glide-slope and lineup control than control pilots not receiving supplemental training. There were also no recycles among the VTRS-trained pilots who transferred to FCLP within two weeks, compared to seven (7.9%) recycles in the control and late transfer groups. There was no transfer advantage for those trained with a daytime high-detail scene. Similarly, there was no transfer advantage for those trained with a wide field of view compared to those trained with a lower cost narrow field-of-view scene. Those pilots who had 40 or 60 simulator trials did considerably better on transfer than the control group, while those who had 20 simulator trials did not. The 60-trial group was not significantly better than the 40-trial group, which suggests a point of diminishing returns after 40 simulator trials. The pilots who trained with a segmented approach schedule did as well or better than those training with either the modified straight-in approach schedule or all-circling approaches. Since the segmented approach schedule involved the least time in the simulator, this method has the advantage and is the recommended method.

REFERENCES

Bickley, W. R. Optimizing simulator-aircraft training mixes. Army Research Institute, Fort Rucker, AL: Proceedings of 2nd Interservice/Industry Training Equipment Conference and Exhibition, National Security Industrial Association, November 1980.

Brictson, C. A. and Burger, W. J. Transfer of training effectiveness: A7E night carrier landing trainer (NCLT), device 2F103. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 74-C-0079-1, 1976.

Brictson, C. A., Burger, W. J., and Wulfeck, J. W. Validation and application of a carrier landing performance score: The LPS. Santa Monica, CA: Dunlap and Associates, Inc., March 1973.

Buckland, G. H., Monroe, E. G., and Mehrer, K. L. Flight simulator runway visual textural cues for landing. Williams AFB, AZ: Operation Screening Division AFHRL-TR-79-81, August 1980.

Cohen, J. Statistical power analysis for the behavioral sciences. New York, NY: Academic Press, 1977.

Collyer, S. C. and Chambers, W. S. AWAVS. a research facility for defining flight trainer visual requirements. Proceedings of the Human Factors Society 22nd Annual Meeting, Santa Monica, CA: Human Factors Society, 1978.

Collyer, S. C., Ricard, G. L., Anderson, M., Westra, D. P., and Perry, R. A. Field of view requirements for carrier landing training. Brooks AFB, TX: Air Force Human Resources Laboratory Report IH-319 AFHRL-TR-80-10, June 1980.

General Electric Co. Systems description--Aviation wide-angle visual system (AWAVS) computer image generator (CIG) visual system. Space Division, Daytona Beach, FL: TR NAVTRAEQUIPCEN 76-C-0048-1, February 1979.

Golovcsenko, I. V. Computer simulation of fresnel lens optical landing system. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN IH-265, September 1976.

Hughes, R. G., Lintern, G., Wightman, D. C., Brooks, R. B., and Singleton, J. Applications of simulator freeze to carrier glideslope tracking instruction. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 78-C-0060-9/ AFHRL-TR-82-3, 1981.

Isley, R. N. and Spears, W. D. Phase I pilot study: VTRS transfer of training experiment. Orlando, L: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN N61339-80-D- 0009-17-2, 1982.

Kraft, C. L., Anderson, C. D., and Elworth, C. L. Peripheral cues and color in visual simulation. Seattle, WA: Data Processing Technology, Boeing Aerospace Company, Final Technical Report, March 1980.

Levine, D. W., and Dunlap, W. P. Power of the F test with skewed data: Should one transform or not? Psychological Bulletin, 1982, 92, 272-280.

Lintern, G. Kaul, C. E., and Collyer, S. C. Glideslope descent-rate cuing to aid carrier landings. Human Factors, 1984, 26, 667-676.

Lintern, G. and Kennedy, R. S. Video game as a covariate for carrier landing research. Perceptual and Motor Skills, 1984, 58, 167-172.

Lintern, G., Mauk, R., and Cotton, J. C. Final approach tracking error measurement for field carrier landing practice. Orlando, FL: Naval Training Systems Center Technical Report NAVTRASYSCEN 85-C-0044-3, in press.

Lintern, G., Sheppard, D. J., McKenna, D., Thomley, K. E., Nolan, M., Wightman, D. C., and Chambers, W. S. Simulator design and instructional features for air-to-ground attack: Transfer study. Orlando, FL: Naval Training Systems Center, NAVTRASYSCEN 85-C-0044-1, 1985.

Lintern, G., Thomley, K., Nelson, B., and Roscoe, S. N. Content, variety, and augmentation of simulated visual scenes for teaching air-to-ground attack. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 81-C- 0105-3, 1984.

McCauley, M. E. and Cotton, J. C. VTRS carrier landing, transfer of training, and performance measurement development. Westlake Village, CA: Canyon Research Group, Inc., March 1982.

Martin, E. L. and Cataneo, D. F. Computer generated image: Relative training effectiveness of day versus night visual scenes. Williams AFB, AZ: Air Force Human Resources Laboratory, AFHRL-TR-79-56, 1980.

Orlansky, J. Current knowledge and projection for assessing the effectiveness of training. Warrenton, VA: Society for Applied Learning Technology. Technology of Training Effectiveness and Evaluation for Productivity Assessment, July 1982.

Povenmire, H. K. and Roscoe, S. N. An evaluation of ground-based flight trainers in routine primary flight training. Human Factors, 1971, 13, 109-116.

Povenmire, H. K. and Roscoe, S. N. Incremental transfer effectiveness of a ground-based general aviation trainer. Human Factors, 1973, 15, 534-542.

Salmoni, A. W., Schmidt, R. A., and Walter, C. B. Knowledge of results and motor learning: A review and critical appraisal. Psychological Bulletin, 1984, 95, 355-386.

Sheppard, D. J. Visual and part-task manipulations for teaching simulated carrier landings. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 81-C-0105-9, 1985.

Simon, C. W. Economical multifactor designs for human factors engineering experiment. Los Angeles, CA: Hughes Aircraft Company Technical Report No. P73-326-A, 1973.

Simon, C. W. New research paradigm for applied experimental psychology: A system approach. Westlake Village, CA: Canyon Research Group, Inc. Technical Report CWS-03-77, 1977.

Singer-Link Division. Aviation wide-angle visual trainer subsystem design report. Binghampton, NY: TR NAVTRAEQUIPCEN 75-C-0009-13, May 1977.

Spears, W. D. Measurement of learning and transfer through curve fitting. Human Factors, 1985, 27, 251-266.

Waag, W. Training effectiveness of visual and motion simulation. Williams AFB, AZ: USAF Human Resources Laboratory, Operations Training Division, AFHRL, USAF, HRL TR-79-72, 1980.

Westra, D. P. Investigation of simulator design features for carrier landing: II. In-simulator transfer of training. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 81-C-0105-1, 1982.

Westra, D. P., and Lintern, G. Simulator design features for helicopter landing on small ships. I. A performance study. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 81-C-0105-13, 1985.

Westra, D. P., Simon, C. W., Collyer, S. C., and Chambers, W. S. Simulator design features for carrier landing: I. Performance experiment. Orlando, FL: Naval Training Equipment Center Technical Report NAVTRAEQUIPCEN 78-C-0060-7, 1981.

Wightman, D. C. Part-task training strategies in simulated carrier landing final approach training. Orlando, FL: Naval Training Equipment Center, Human Factors Laboratory Technical Report NAVTRAEQUIPCEN IH-347, 1983.

Wightman, D. C., Westra, D. P., and Lintern, G. Transfer of training of simulator visual and training features for the carrier landing task with undergraduate pilots. Proceedings of the 7th Interservice/Industry Training Equipment Conference, Orlando, FL: 1985.

APPENDIX A

FRESNEL LENS OPTICAL LANDING SYSTEM

The Fresnel Lens Optical Landing System (FLOLS) provides primary glideslope displacement information for a carrier approach to landing. It consists of light sources behind five vertically-stacked Fresnel lenses that are situated between two horizontal light arrays known as the datum bars. The array of lenses and lamps provides a virtual image which appears to the pilot as a single light located 150 feet behind the datum bars. This light is known as the meatball. It is visible to the pilot through the center lens and is seen as level with the datum bars when the aircraft is within 9.5 minutes of arc of the glideslope. As the aircraft moves more than 9.5 minutes of arc above or below the glideslope, the meatball is seen through higher or lower Fresnel lenses to give the appearance of moving vertically above or below the line of the datum bars Figure A-1 depicts the FLOLS and its projection aft of the carrier.

For a carrier approach the pilot attempts to follow a designated glideslope (usually 3.5 degrees) by keeping the meatball level with the datum bars, so that a hook attached to the tail of the aircraft will contact the landing deck midway between the second and third of four arrestment cables. These cables (more frequently referred to as wires) are stretched across the landing deck at different distances from the ramp. Under the aircraft's momentum the hook travels forward to snag the third wire for a trap (arrested landing). The first or second wire may be caught on a low approach and the fourth on a high approach. Very low approaches can result in a ramp strike while high approaches can result in a bolter (a missed approach because of touchdown beyond the wire arrestment area).

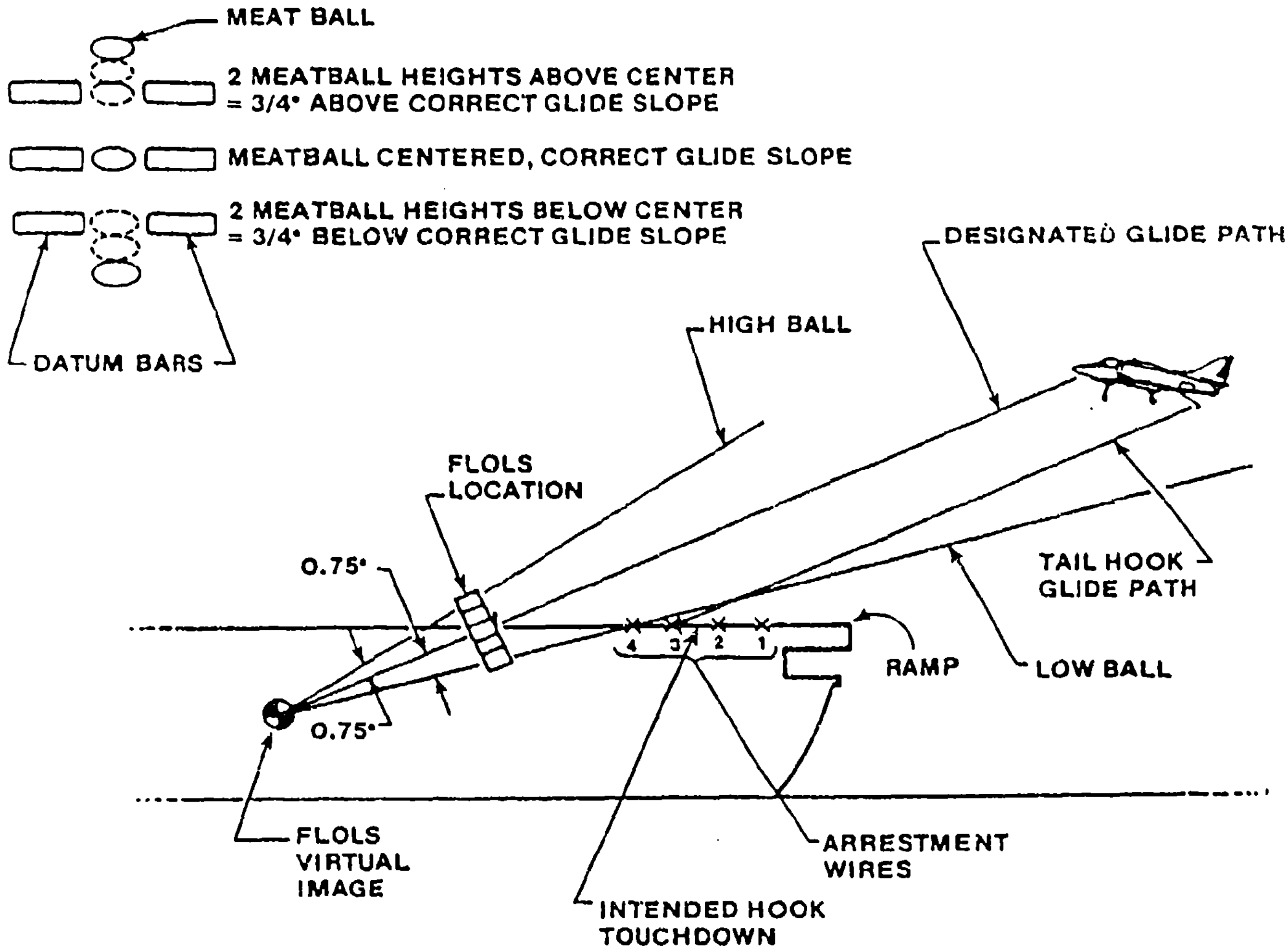

Figure A-1. Carrier Approach Geometry Depicting FLOLS Projection of Glideslope Deviation Information. Adapted from Golovcsenko (1976).

APPENDIX B

INSTRUCTOR AIDS

The Instructor/Operator (IO) station provides the capability of interacting with the computer and flight simulator and monitoring pilot performance. Among the principal features of the IO station are six monitors. Two alphanumeric monitors are used by the operator to interact with the computer and flight simulator to develop, control, and record the experiment. Two color video monitors display the images from the two channels of the computer image generator (CIG) and two graphic monitors present graphic displays of the operator or instructor's choice (Figure B-1).

The video and graphic monitors were used by the instructors to monitor pilot's performance during the experiment. The video monitors gave a general perspective of what the pilot was seeing in the simulator. One monitor displayed the background scene (seascape and sky) and the other monitor displayed the high resolution target image (carrier). The two graphic monitors provided the LSO with feedback on the student pilot's performance. One display was a real-time view of major cockpit instruments (Format C display) and the other display presented a time history and performance measure plotted from a distance of 6000 feet from the carrier to touchdown (LSO display).

The layout of Format C is shown in Figure B-2. Most of the instruments displayed in Format C were replications of the T2-C cockpit instruments. Thus, the instructor could monitor the pilot's control inputs and the aircraft's flight characteristics and position for each pass. A list of the instruments displayed in Format C are shown in Table B-1.

The layout of the LSO display is shown in Figure B-3. The display provided the LSO with a time history of each carrier approach for six performance and control measures. The parameters were plotted along the horizontal axis of the LSO display with the solid lines indicating optimum performance or control settings, and the dotted lines representing actual performance. These measures are described in Table B-2.

TECHNICAL INFORMATION

The computer system consists of multiple SEL 32/77 central processing units. The graphic monitors were manufactured by Vector General. The software to generate the Format C display was written and cataloged under the SEL assembler. Two modules comprised the Format C software program. The Vector General assembly display list software (30 pages) generated the skeleton (nondynamic) display of the main instrumentation of

Format C. The Vector General Buffer generated the static display from this program.

The dynamic changes of Format C were driven by the SEL assembly drive module (25 pages). This program was the driver routine for the Format C display. On the first pass-through this routine the graphics status word is reset and the Vector General code is read from the disc into the Vector General Buffer. This routine then updates the values used to draw the current position and values of the instruments. The LSO display was generated in a similar manner as the Format C display. A copy of the software to generate these displays is available upon request by contacting the Visual Technology Research Simulator.

Two Alphanumeric Monitors (Top Center, Bottom Center)
Two Color Video Monitors (Top Left, Top Right)
Two Graphic Monitors (Bottom Left, Bottom Right)

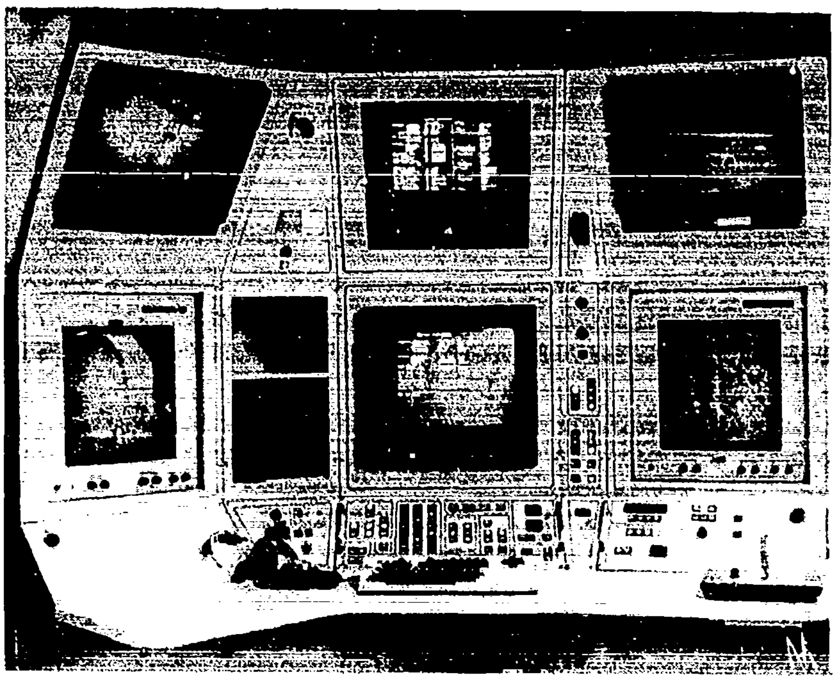

Figure B-1. Instructor/Operator Station

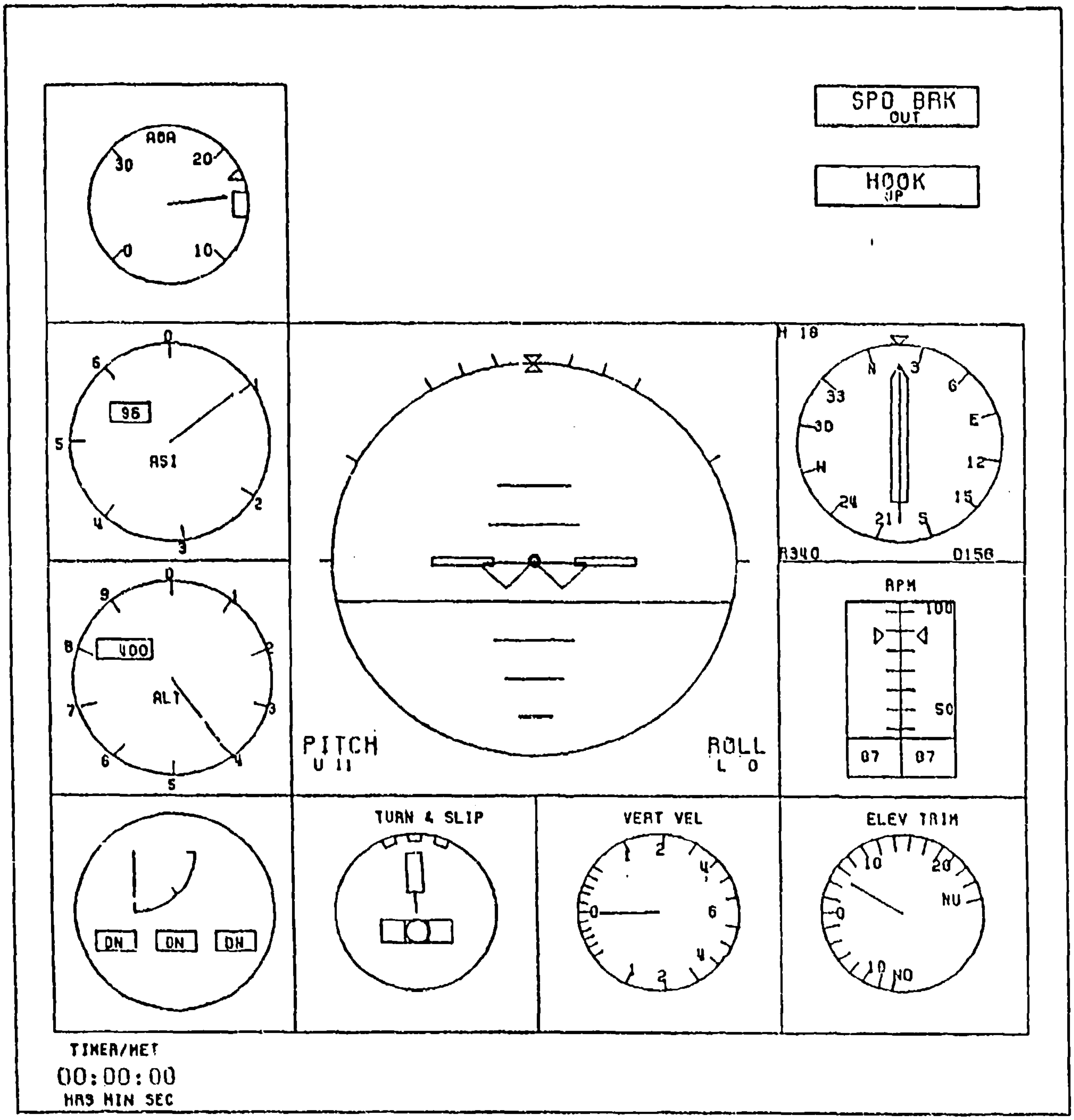

Figure B-2. Layout of Format C. Instruments Shown are Identified in Table B-1.

Table B-1. INSTRUMENTS DISPLAYED IN FORMAT C

Angle of Attack Indicator (AOA)
Airspeed Indicator (ASI)
Altitude Indicator (ALT)
Attitude Indicator
Power Gauge (RPM)
Elevator Trim Indicator (Elev Trim)
Vertical Speed Indicator (Vert Vel)
Turn and Slip Indicator
Flaps Indicator
Speed Break Indicator (Spd Brk)
Hook Indicator

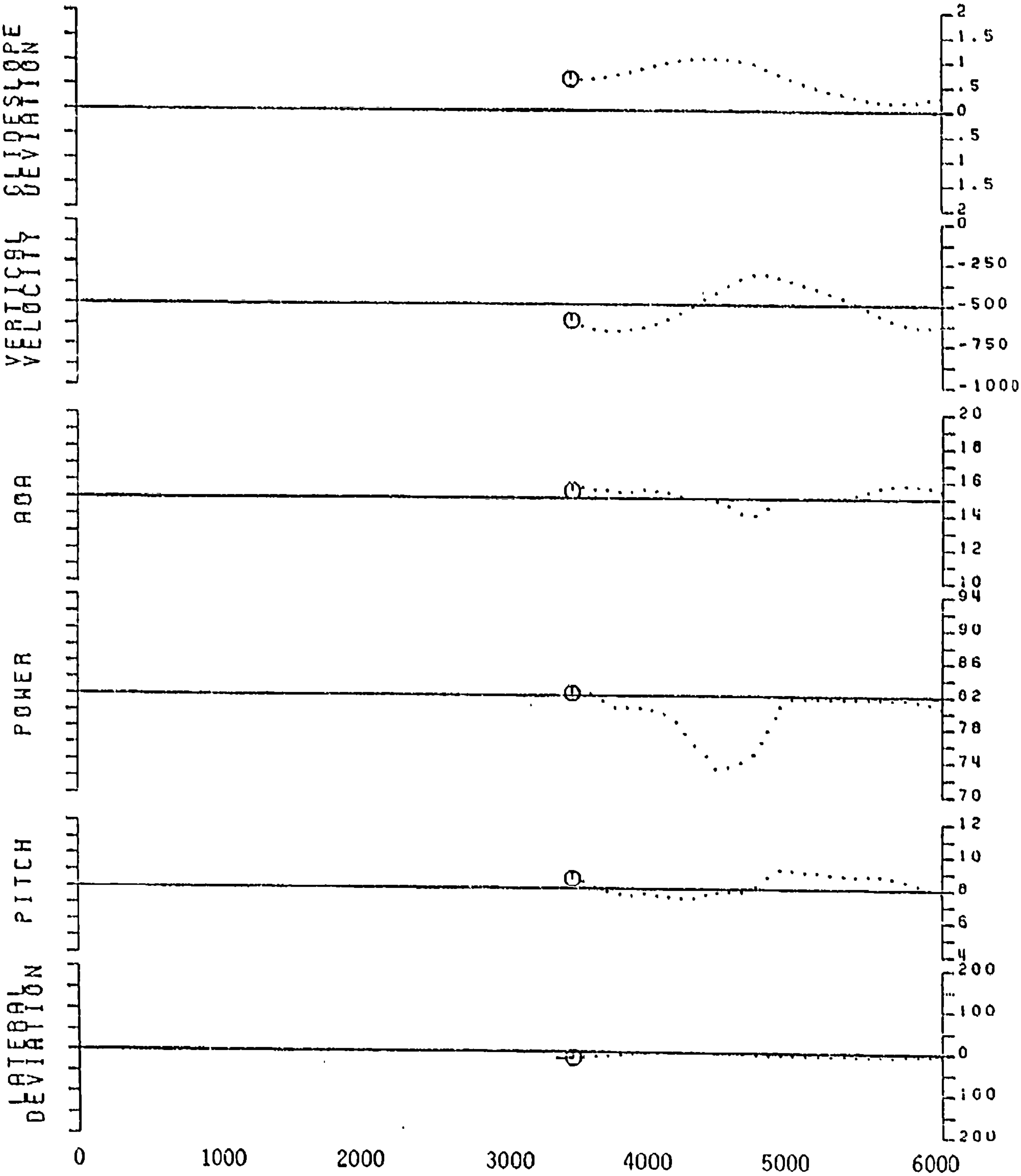

Figure B-3. Layout of LSO Display. Performance Indices are Summarized in Table B-2.

TABLE B-2. PERFORMANCE MEASURES PLOTTED IN THE LSO DISPLAY

Performance Measures	Optimum	Range of Deviation
Glideslope Deviation	On glideslope (Center Ball)	Between two balls above and below glideslope
Lateral Deviation	On centerline	Between 200 ft left and right of centerline
Angle of Attack (AOA)	15 Units AOA	Between 10 and 20 units AOA
Vertical Velocity	500 RPM rate of descent	Between 0 and 1000 FPM rate of descent
Power	82%	Between 70% and 94%
Pitch	Eight degrees pitch up	Between 4 and 12 degrees pitch up

APPENDIX C

FIELD PERFORMANCE MEASUREMENT

A laser tracking system, known as HYTAL for Hybrid Terminal Assist Landing, was selected for field data acquisition. It is a brass-board, laboratory demonstration system that was developed at Naval Weapons Center, China Lake, as a backup for the microwave automatic carrier landing system. Many factors, including cost, availability, and potential accuracy entered into the selection of this device for our study (McCauley and Cotton, 1982).

The HYTAL system was placed near the active runway (Figure C-1) and used to measure vertical and lateral deviations from the approach glidepath during field carrier landing practice. This was accomplished with a laser transmitter/receiver tracking head (Figure C-2) that rotated in azimuth and elevation. A low-power laser signal (eye-safe at 4 feet) was transmitted towards the approaching aircraft as it turned onto the final approach after the tracking head was manually aligned in its general direction. Each of the test aircraft was fitted with a glass retroreflector to return the laser signal (Figure C-3). The tracking head searched in the vicinity of the aircraft until it found and locked onto the return signal. The tracking head remained locked onto the direction of maximum signal return until the aircraft touched down on the runway and passed the tracking station.

Aircraft position was sampled at 20 Hz. Range was determined by the time required for a laser pulse to travel to the aircraft and back to the laser receiver. Azimuth and elevation angles of the HYTAL tracking head were transformed into electrical digitized signals which were used to determine azimuth and elevation angles of the approaching aircraft. The necessary transformations of the electronic signals to readouts of range, azimuth, and elevation were accomplished with integrated circuits. The final readings were stored on audio track number 1 of a video tape. The video channel was used to store a visual record of the aircraft approach, and audio track number 2 for a record of LSO comments and wave-off instructions. Wind speed and direction were also recorded for each approach.

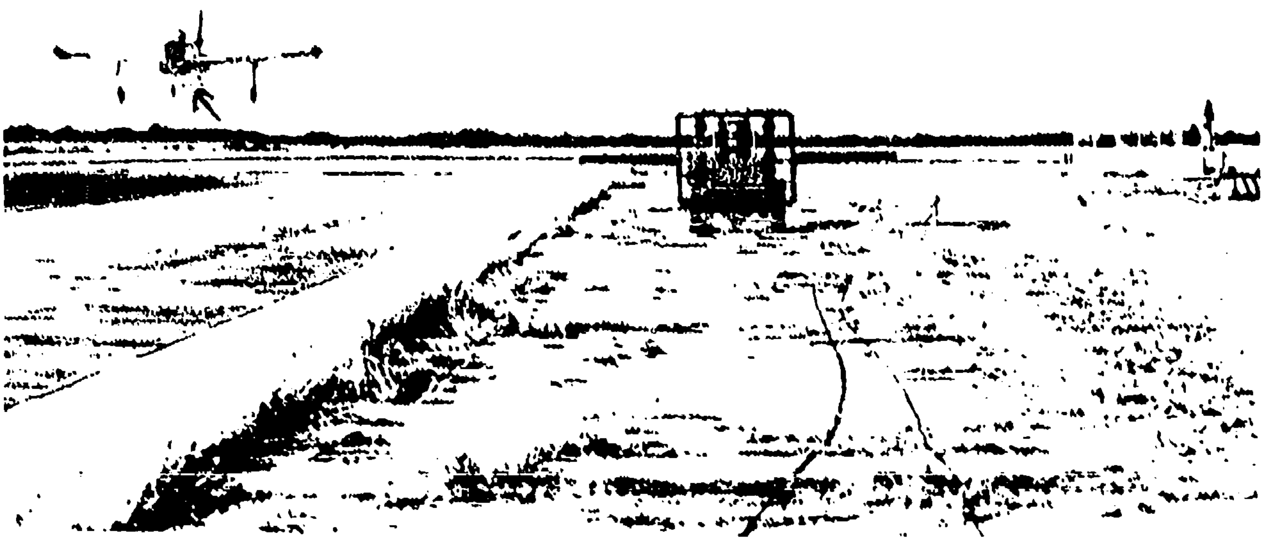

Figure C-1. A T-2C About to Track Down During FCLP. The Retroreflector is Indicated by the Arrow. The rear of the FLOLS is at Center Foreground. The LSO Shack is to the Right. The T-2C in the Background is Turning Onto Final and is Approximately 30 Seconds from Touchdown.

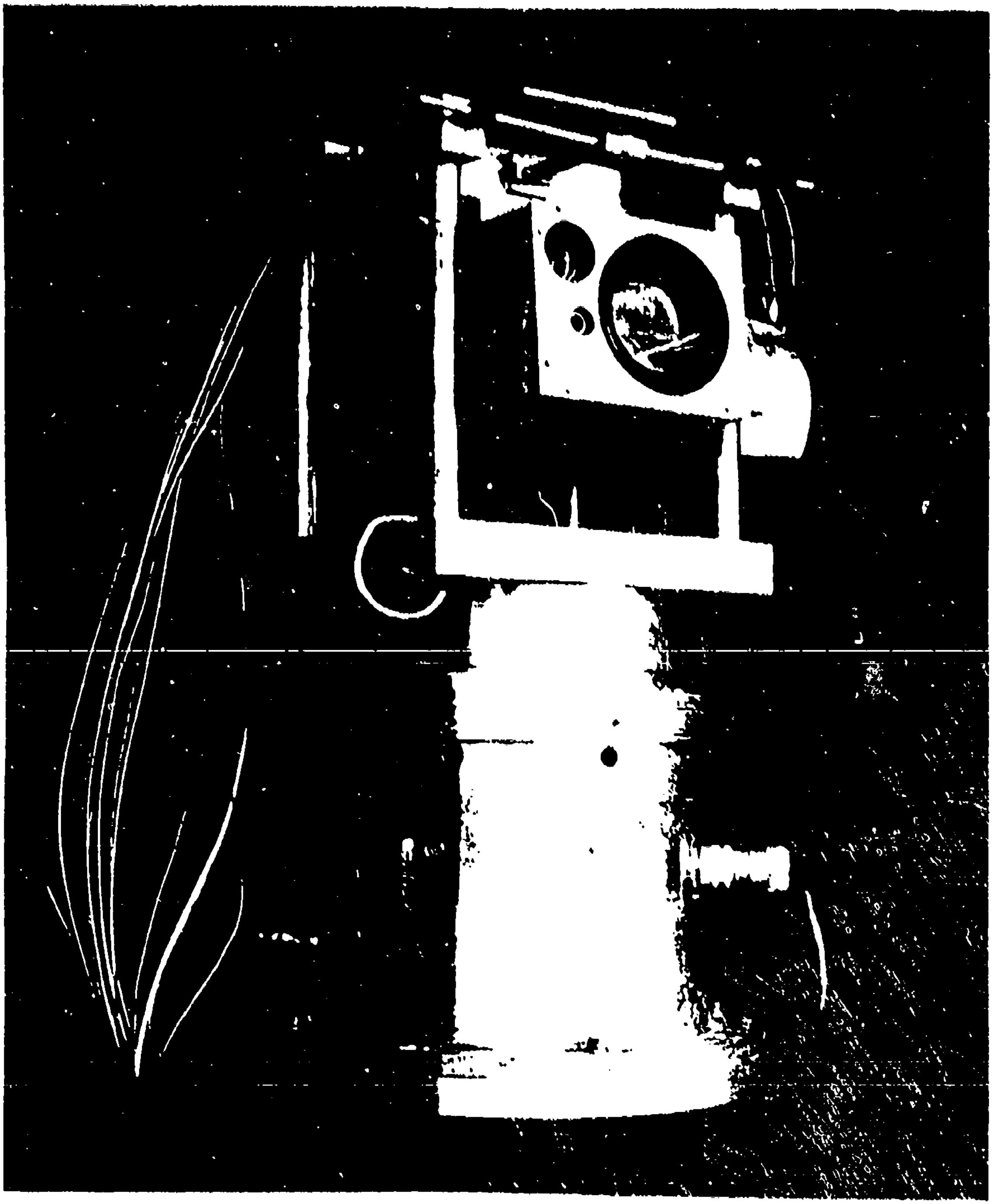

Figure C-2. HYTAL Tracking Head as Mounted in the Van Ready for Glideslope Tracking

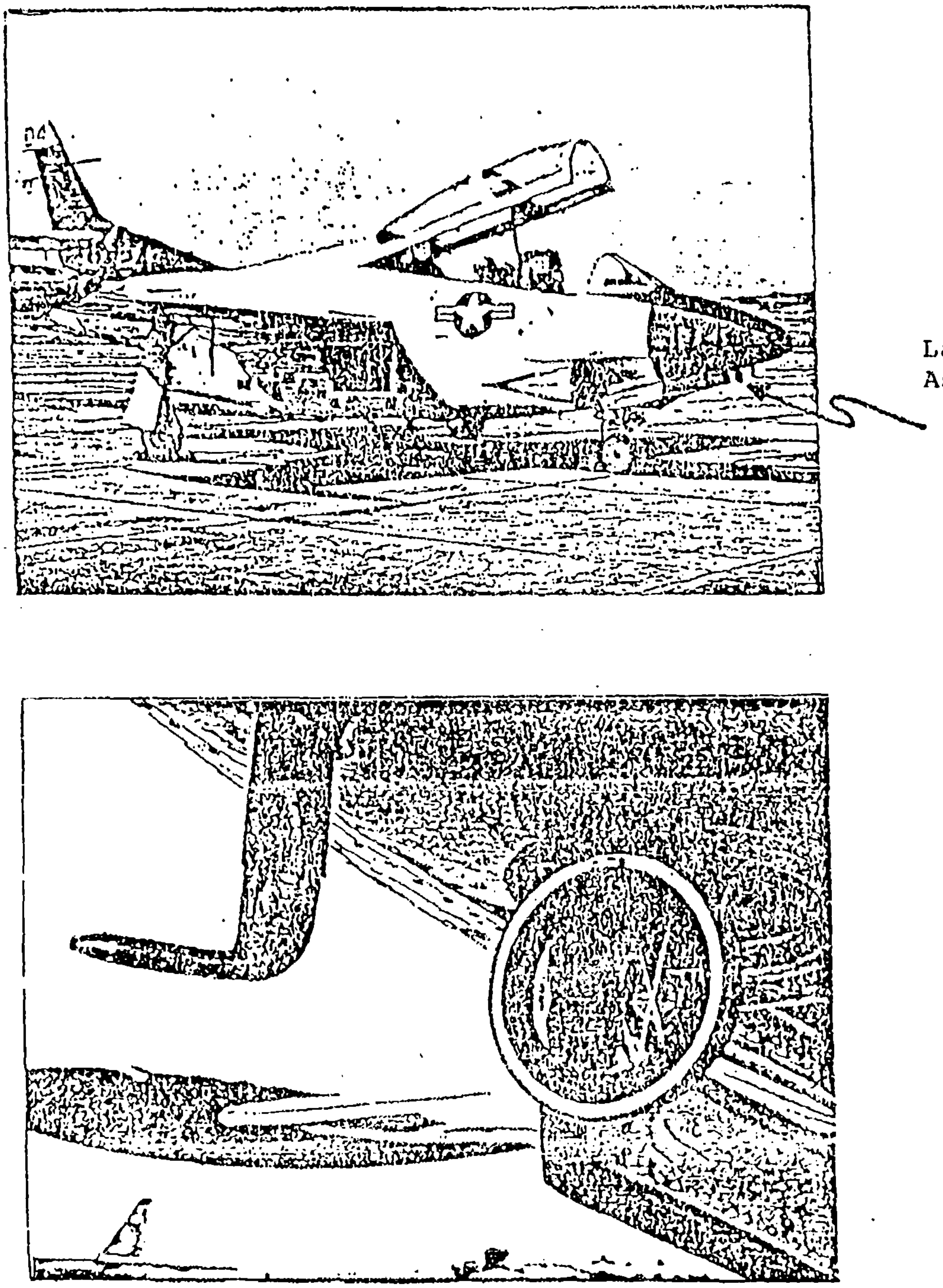

Figure C-3. Two Views of the Retroreflector
Mounted in a T-2C Landing Light Housing

HYTAL azimuth and elevation accuracy is rated at +/- 0.1 degrees (+/- 6 arc minutes). Range accuracy is rated at +/- 2 feet at one mile. Small errors in some of the constants used to calculate aircraft elevation and azimuth were presumably introduced by the need to move the HYTAL system occasionally. The equipment was calibrated before the commencement of trials each morning. Wind shifts occasionally forced a change in the active runway and it was necessary to move the HYTAL tracking system at these times. Calibrations were checked each evening after the completion of the day's flight trials. Range and elevation were calibrated with a retroreflector mounted on a 5-foot pole which was placed vertically on a concrete pad 230 feet from the HYTAL tracking head. For the elevation calibration, information about relative ground height at the concrete pad and HYTAL head was essential. This was measured by a water-level method as illustrated in Figure C-4. Azimuth was calibrated with the calibration pole placed vertically at the nominal touchdown point on the runway. Discrepancies between actual and measured range, elevation, and azimuth at calibration were used to adjust the data prior to statistical analysis.

Proper calibration of the FLOLS was also considered essential. The nominal glideslope was 3.25 degrees. Any inadvertent error in the glideslope angle indicated by the FLOLS would induce pilots to fly below or above the HYTAL "zero-error" glideslope. LSOs calibrated the FLOLS with an angled mirror placed on a pole approximately 18 feet in length. The calibration spot was a surveyed, concrete pad, 150 feet in front of the FLOLS (Figure C-5). The length of the pole can be adjusted for various FLOLS settings. Under normal calibration procedures, the LSO adjusts the pole to the correct setting and then stands it on the concrete calibration point. The mirror is then at the height that a pilot's eye would be if flying down the glideslope and at that distance from the FLOLS. The LSO should see a center ball in the FLOLS by looking into the angled mirror. Experimenters noted that LSOs would accept slight but noticeable deviations from center ball. Thus, the experimenters checked the FLOLS calibration each day before the LSOs arrived at the site and advised them if they thought adjustments were desirable. Experimenters judged that accuracy of the FLOLS setting was maintained within +/- 0.08 degrees by this procedure.

System calibration was further checked with one series of test flights. An LSO flew approaches in a test aircraft with intentional deviations to the left or the right, and above or below the glideslope. Inspection of data from these trials showed deviations similar to those intended by the LSO. In particular, directions of errors were carefully checked.

Digital data were lifted from the video tape and converted by hardware to VAX compatible format. During this process,

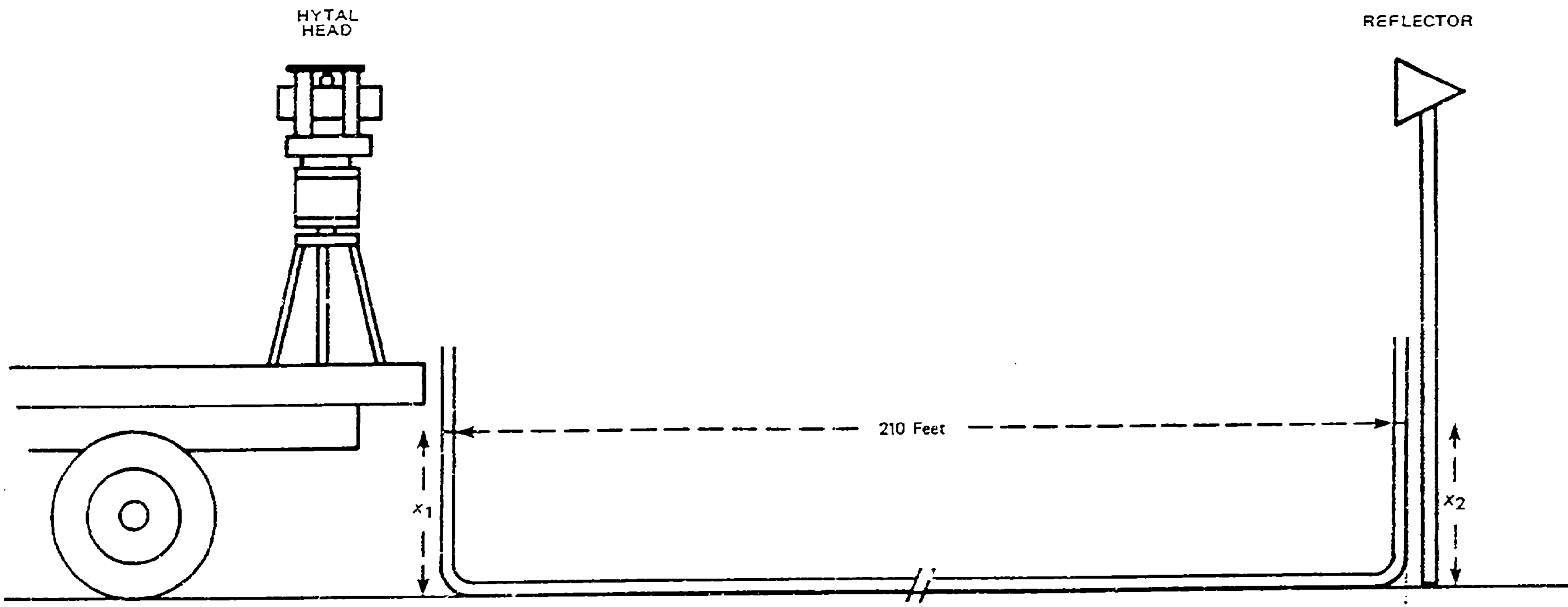

Figure C-4. Measurement of Relative Ground Height at Test Reflector and HYTAL Head by the Water Level Method. (The water level at each end of the tube is at the same height. Any discrepancy between x_1 and x_2 indicates a difference in ground height at the two positions.)

Figure C-5. FLOLS Calibration Procedure Used by Landing Signal Officers to Set the FLOLS Glideslope subjects, trials, and conditions were identified from VTRS data listings, from LSO field records, and from information on audio tracks 1 and 2. Care was taken by integrating information from various sources to ensure that trials had been identified correctly. This care was essential because aircraft often turned to final at 30-second intervals and would be tracked for approximately 20 to 30 seconds. Occasionally, some of the aircraft in the patterns were not involved in the experiment. Trial identification information was included in the computer records as header information.

The "bang-bang" operation of the laser tracking head produced some instability in the last bit of the digital record of stored values. This caused the azimuth and elevation data to dither by approximately 0.5 degrees. A moving filter was applied to the raw data to smooth this instability. Figure C-6 is a representation of the filter mechanization used for this purpose. At the start of the data stream, outputting of the first average was suppressed until the seventh value was received. Subsequently, the averaging "gate" moved along one value at a time to provide the average for the last seven values of the raw data. Accordingly, the filter had a time constant of approximately one-third of a second which is considered inconsequential for practical measurement purposes. At the same time, the data storage rate was reduced from 20 Hz to 5 Hz, that being considered adequate for any conceivable analyses.

Trigonometric transformations programmed in software were used to convert the HYTAL-referenced values to angles and distances that were referenced to the nominal glideslope and nominal point of touchdown.

The horizontal and vertical geometry is shown in Figures C-7 and C-8. The following constants are used in the calculations.

VRR = Vertical distance between the retroreflector and the bottom of the main wheels with the oleo strut extended.

LRR = Lateral distance between the retroreflector and the vertical axis of the main wheels.

VHD = Vertical distance between the HYTAL sensing head horizontal reference line and the center of the FCLP deck.

HSC = Lateral distance between the HYTAL sensing head lateral reference line and the center line of the FCLP deck.

LTD = Lateral distance between the HYTAL sensing head and the nominal touchdown point.

The following output variables were provided by the HYTAL sensing device.

SRR = Slant range from the HYTAL sensing head to the retroreflector on the nose of the aircraft.

ΘHS = Azimuth angle to the aircraft retroreflector relative to the HYTAL sensing head vertical reference line.

ΦHS = Elevation angle to the aircraft retroreflector relative to the HYTAL sensing head lateral reference line.

The instantaneous laterally corrected range was calculated from:

$$SCR = SRR, \cos(\Phi HS) \cos(\Theta HS),$$

and the instantaneous corrected range from the aircraft wheels to the touchtown point from:

$$RTD = SCR + LRR - LTD$$

and the instantaneous lateral displacement from the HYTAL lateral reference line from:

$$HLD = SRR \sin(\Phi HS) \cos(\Theta HS)$$

Therefore, the aircraft's displacement from the FCLP deck centerline at any instant was:

$$HAP = HLD - HSC \text{ (for given left/right convention).}$$

The instantaneous altitude of the aircraft altitude above the HYTAL horizontal reference line was:

$$VHA = SRR \sin (\Theta HS)$$

and the instantaneous corrected altitude above HYTAL horizontal reference line was:

$$VHC = VHA - VRR$$

Therefore, the instantaneous corrected altitude above the FCLP deck center was:

$$VAC = VHC + VHD$$

Test wave-offs (one was generally given for each student pilot in each event) were eliminated from consideration.

Performance wave-offs were retained in the data set, but tracking scores were not considered beyond the point at which the LSO gave the wave-off instruction.

The only modification of LSO or student pilot duties in the field trials related to the use of the retroreflectors. These were mounted in the landing light housing of the T-2C aircraft (Figure C-3) so that students were required to lower the landing light assembly before they turned to final for their first approach. As the landing lights are not normally lowered for day FCLP, LSOs reminded their students of this requirement when they were flying their first landing pattern. An experimenter checked that the retroreflectors were lowered on first approaches, and requested LSOs to remind student pilots on those occasions they were not.

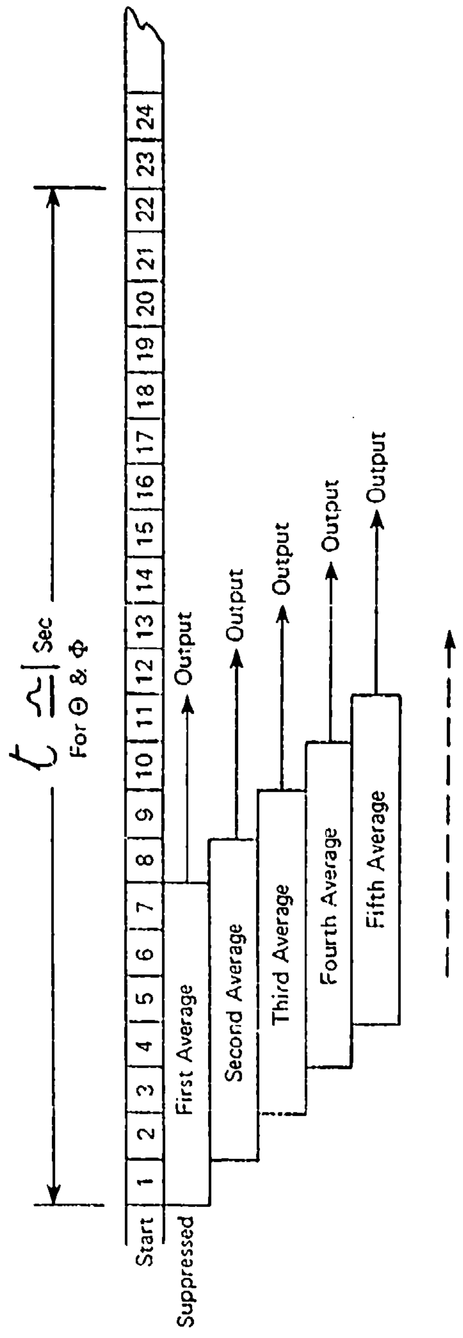

Figure C-6. Moving Average Filter Mechanization.

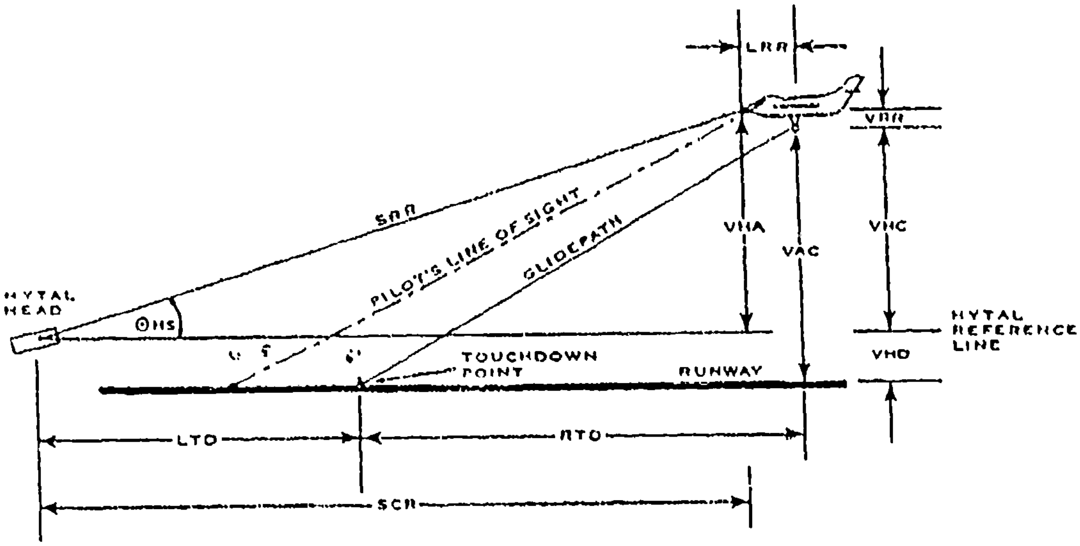

Figure C-7. Calculation of Height of Main Wheels Above Touchdown

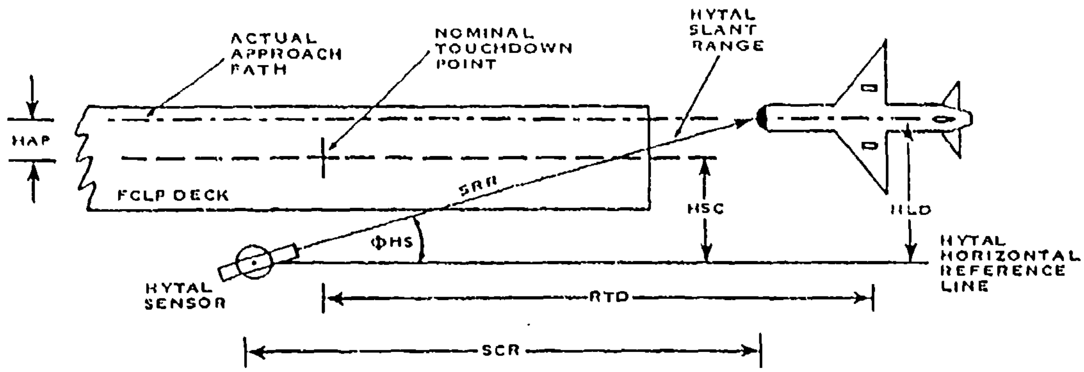

Figure C-8. Top View of Range of Retroreflector from the HYTAL Head